OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:14:12 ; Search time 158.02 Seconds
                 (without alignments)
                 5.156 Million cell updates/sec

Title: us-09-281-760c-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

● minimum DB seq length: 0
Maximum DB seq length: 2000000000

Total number of hits satisfying chosen parameters: 522463

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   A_Geneseq_1101:*
                 1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                 2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                 3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                 4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                 5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                 6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                 7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                 8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                 9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	63	100.0	11	21	AAY69585	Canine IgE-derived
2	63	100.0	11	21	AAY50879	8H.8 binding canin
3	63	100.0	20	18	AAW24098	Canine immunoglobi
4	63	100.0	124	18	AAW24097	Partial canine imm
5	63	100.0	312	21	AAY79995	Dog Immunoglobulin
6	63	100.0	341	21	AAB06208	Immunogenic peptid
7	63	100.0	417	18	AAW23067	Canine IgE heavy c
8	63	100.0	426	17	AAR97753	Canine IgE. Canis
9	63	93.7	22	21	AAY69587	Peptide E3A.5-exte
10	54	85.7	11	21	AAY50882	Antibody 8H.8 bind
11	43.5	69.0	12	21	AAY69592	Canine substituted

RESULT   1
ID   AAY69585 standard; peptide; 11 AA.
XX
AC   AAY69585;
XX
DT   19-APR-2000 (first entry)
XX
DE   Canine IgE-derived peptide E3a.5, SEQ ID NO:10.
XX
KW   Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW   allergy; canine; prophylaxis; treatment.
XX
OS   Canis familiaris.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107020.
XX
PR   09-APR-1998; 98US-0058332.
PR   30-MAR-1999; 99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.

ALIGNMENTS

12	13	41	351	13	AAR24251	Vaccinia virus B18
13	14	41	351	14	AAR41542	B15R product. Vac
14	21	41	574	21	AAY33969	Amino acid sequenc
15	22	41	582	22	AAB78858	Human lung tumour
16	17	41	583	17	AAW06891	Human activated le
17	19	41	583	19	AAR97230	Stem cell marker H
18	17	41	583	17	AAW47088	Activated leukocyt
19	21	40	11	21	AAY59591	Feline IgE-derived
20	22	40	11	22	AAY50881	Antibody 8H.8 bind
21	21	63.5	343	21	AAB06204	Platypus IgE heavy
22	22	60.3	107	22	AAG93603	Human anti-Rh(D) c
23	20	60.3	130	20	AAW97813	Bovine butyrophili
24	22	60.3	193	22	AAB60775	Gene 47 related pe
25	18	60.3	196	18	AAW10505	Alphalalpha2 regio
26	18	60.3	196	18	AAW10508	Alphalalpha2 regio
27	18	60.3	234	18	AAW10509	Soluble fused MHC
28	21	60.3	256	21	AAY68281	Class II alpha cha
29	21	60.3	256	21	AAY52935	Class II alpha-cha
30	22	60.3	256	22	AAB86896	Class II alpha cha
31	18	60.3	441	18	AAW10507	Soluble fused MHC
32	18	60.3	448	18	AAW10513	Soluble fused MHC
33	17	60.3	458	17	AAR98907	Vector SCEI-derive
34	18	60.3	458	18	AAW22214	SCEI single chain
35	18	60.3	459	18	AAR98905	Vector SSCI-derive
36	18	60.3	459	18	AAW29212	SSC1 single chain
37	22	60.3	494	22	AAB67480	A fusion protein c
38	18	60.3	500	18	AAR98906	Vector SCR1-derive
39	17	60.3	500	17	AAW29213	SCR1 single chain
40	20	60.3	500	20	AAY27111	Single chain 1A0/o
41	20	60.3	526	20	AAW97812	Bovine butyrophili
42	22	60.3	26926	22	AAU05396	Human titin (conne
43	14	58.7	860	14	AAR42962	Bovine MTP.
44	14	58.7	894	14	AAR42963	Human MTP. Bos t
45	17	58.7	894	17	AAR50007	Human MTP. Homo s
46	15	57.1	96	15	AAR54323	Anti-HIV gp120 imm
47	17	57.1	96	17	AAW01292	VL region of HIV n
48	21	57.1	96	21	AAY95144	Anti-gp120 antibod
49	21	57.1	96	21	AAY98253	Anti-gp120 antibod
50	20	57.1	108	20	AAY12352	Human 5' EST secre

```
XX
XX
XX
PT   New peptides, useful for generating canine autoantibodies against
PS   immunoglobulin E (IgE) to treat allergic symptoms -
XX
     Claim 5; Page 5; 30pp; English.
XX
CC   Sequences AAY69576-Y69586 and AAY69598-Y69599 represent peptides that may
CC   be used to induce anti-IgE (immunoglobulin E) autoantibodies in a dog.
CC   IgE is an important mediator of canine allergic responses, including Type
CC   I immediate hypersensitivity. It is estimated that up to 30% of all dogs
CC   suffer from allergies or allergy-related skin disorders, and allergic
CC   dermatitis is thought to affect between 3 and 15% of dogs. Common canine
CC   allergens include fleas, pollens, moulds and dust, with flea allergy
CC   being the most frequent. Once IgE has been released in response to
CC   allergen, the IgE binds to receptors on mast cells. The IgE molecules
CC   can then become cross-linked, which induces the mast cells to release
CC   histamine, which leads to the manifestation of allergic symptoms. The
CC   peptides of the invention correspond to epitopes of canine IgE, or
CC   conservative variations thereof, that are accessible in circulating IgE
CC   but not in IgE bound to mast cells. The autoantibodies produced on
CC   immunisation with such peptides therefore target only circulating IgE,
CC   avoiding the possibility that they may cross-link bound IgE molecules,
CC   which would exacerbate an ongoing allergic response even in the absence
CC   of allergen. Anti-IgE autoantibody induction results in an immune
CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targeting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog.
XX
SQ   Sequence   11 AA;

Query Match             100.0%;  Score 63;  DB 21;  Length 11;
Best Local Similarity   100.0%;  Pred. No. 7.9e-05;
Matches  11; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GMNLTWYRESK 11
              |||||||||||
Db          1 gmnltwyresk 11

RESULT 2
ID   AAY50879
     AAY50879 standard; peptide; 11 AA.
AC   AAY50879;
DT   24-FEB-2000  (first entry)
DE   8H.8 binding canine mimotope 1.
KW   Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW   epitope; prophylaxis; treatment; mimotope.
OS   Synthetic.
PN   EP957111-A2.
PD   17-NOV-1999.
PF   09-APR-1999;  99EP-0107035.
PR   09-APR-1998;  98US-0058331.
PR   30-MAR-1999;  99US-0281760.
PA   (IDEX-) IDEXX LAB INC.
PI   Lawton R, Mermer B, Francoeur G;
DR   WPI; 2000-040833/04.
PT   Binding proteins used for treatment or prophylaxis of canine allergy -

XX
PS   Claim 10; Page 13; 30pp; English.
XX
CC   This invention describes a novel binding protein which specifically
CC   binds to native canine free or B-cell bound IgE, and which doesn't bind
CC   to IgE when the IgE is bound to mast cells. The peptide products of the
CC   invention have anti-IgE activity. The antibodies bind to defined
CC   epitopes on free or B-cell bound IgE molecules which have an important
CC   role in allergic reaction. The specific binding proteins are used to
CC   produce a pharmaceutical composition, preferably with a diluent, which
CC   can be used for prophylaxis or treatment of canine allergy.
CC   AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC   invention.
XX
SQ   Sequence   11 AA;

Query Match             100.0%;  Score 63;  DB 21;  Length 11;
Best Local Similarity   100.0%;  Pred. No. 7.9e-05;
Matches  11; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GMNLTWYRESK 11
              |||||||||||
Db          1 gmnltwyresk 11

RESULT 3
ID   AAW24098
     AAW24098 standard; peptide; 20 AA.
AC   AAW24098;
DT   21-NOV-1997  (first entry)
DE   Canine immunoglobin E peptide 1.
KW   Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
OS   Canis familiaris.
PN   JP09169795-A.
PD   30-JUN-1997.
PF   22-DEC-1995;  95JP-0334381.
PR   22-DEC-1995;  95JP-0334381.
PA   (HITB ) HITACHI CHEM CO LTD.
DR   WPI; 1997-389423/36.
DR   N-PSDB; AAW85647.
PT   Canine immunoglobulin E peptide fragment and related DNA - useful
PT   for the preparation of anti-canine immunoglobulin E antibody
XX
PS   Claim 2; Page 9; 12pp; Japanese.
XX
CC   Canine immunoglobulin E (IgE) peptide fragments containing at least 5 continuous
CC   amino acids of the partial canine immunoglobulin E (IgE) protein shown
CC   in AAW24097. The peptides are used for the preparation of anti-canine
CC   IgE antibody. The anti-canine IgE antibody can be used for the diagnosis
CC   of canine allergies.
XX
SQ   Sequence   20 AA;

Query Match             100.0%;  Score 63;  DB 18;  Length 20;
Best Local Similarity   100.0%;  Pred. No. 0.00015;
Matches  11; Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 GMNLTWYRESK 11
              |||||||||||
```

Db   3 gmnltwyresk 13

RESULT 4
ID   AAW24097 standard; peptide; 124 AA.
XX
AC   AAW24097;
XX
DT   21-NOV-1997 (first entry)
XX
DE   Partial canine immunoglobin E protein.
XX
KW   Immunoglobulin E; IgE; anti-canine IgE antibody; allergy; canine; dog.
XX
OS   Canis familiaris.
XX
PN   JP09169995-A;
XX
PD   30-JUN-1997.
XX
PF   22-DEC-1995;   95JP-0334381.
XX
PR   22-DEC-1995;   95JP-0334381.
XX
PA   (HITB) HITACHI CHEM CO LTD.
XX
DR   WPI; 1997-389423/36.
DR   N-PSDB; AAT85646.
XX
PT   Canine immunoglobulin E peptide fragment and related DNA - useful
PT   for the preparation of anti-canine immunoglobulin E antibody
XX
PS   Claim 1; Page 8; 12pp; Japanese.
XX
CC   This is a partial canine immunoglobulin E (IgE) protein. Peptide
CC   fragments (AAW24098-106) containing at least five continuous amino acids
CC   of this sequence are used for the preparation of anti-canine IgE
CC   antibody. The anti-canine IgE antibody can be used for the diagnosis of
CC   canine allergies.
XX
SQ   Sequence   124 AA;

Qy   1 GMNLTWYRESK 11
       |||||||||||
Db   13 gmnltwyresk 23

Query Match   100.0%; Score 63; DB 18; Length 124;
Best Local Similarity 100.0%; Pred. No. 0.0011;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 5
ID   AAY79995 standard; protein; 312 AA.
XX
AC   AAY79995;
XX
DT   15-MAY-2000 (first entry)
XX
DE   Dog immunoglobulin E epsilon heavy chain SEQ ID NO:2.
XX
KW   Immunoglobulin E; IgE; epsilon heavy chain; antigenic; antigen;
KW   immunogenic; immunostimulatory; carrier protein; helper T cell epitope;
KW   antibody; allergy; allergic disease; immunisation; anti-allergic;
KW   anti-anaphylactic; anti-asthmatic; asthma; anaphylaxis; dermatitis.
XX
OS   Canis sp.
XX
PN   WO9967293-A1;
XX
PD   29-DEC-1999.

XX
PF   21-JUN-1999;   99WO-US13959.
XX
PR   20-JUN-1998;   98US-0100287.
XX
PA   (UNBI-) UNITED BIOMEDICAL INC.
XX
PA   Wang CY, Walfield AM;
XX
DR   WPI; 2000-160578/14.
XX
PT   New antigenic peptide from the CH3 domain of immunoglobulin E, fusions
PT   for immunization against allergy -
XX
CC   Example 1; Page 66-68; 155pp; English.
XX
CC   The present invention describes immunoglobulin E (IgE)-CH3 domain
CC   antigenic peptides (I). (II) have anti-allergic, anti-anaphylactic
CC   and anti-asthmatic properties. (I) induces polyclonal antibodies
CC   specific for a target effector site on the epsilon-heavy chain of IgE,
CC   and so preventing triggering and activation of mast cells and basophils
CC   and downregulation of IgE synthesis. Conjugates, or fusion peptides,
CC   containing (I) are used for active immunisation against IgE-mediated
CC   allergies, e.g. food allergies, asthma, anaphylaxis, or flea-allergy
CC   dermatitis. Nucleic acids that encode these compounds are useful for
CC   recombinant production of corresponding peptides or in DNA vaccines.
CC   Conjugates of (I) that include a promiscuous T helper cell epitope
CC   (functional in genetically diverse subjects), in addition to a B cell
CC   target epitope, have increased immunogenicity and may include cyclic
CC   constraints (disulfide bridge) to stabilise conformational features and
CC   maximize cross-reactivity to the natural target. They induce safe
CC   (non-anaphylactogenic) antibodies. AAY79994 to AAY80084 represent amino
CC   acid sequences used in the exemplification of the present invention.
XX
SQ   Sequence   312 AA;

Qy   1 GMNLTWYRESK 11
       |||||||||||
Db   143 gmnltwyresk 153

Query Match   100.0%; Score 63; DB 21; Length 312;
Best Local Similarity 100.0%; Pred. No. 0.0031;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 6
ID   AAB06208 standard; protein; 341 AA.
XX
AC   AAB06208;
XX
DT   22-NOV-2000 (first entry)
XX
DE   Immunogenic peptide consisting of opossum CH2, dog CH3 and opossum CH4.
XX
KW   Dog; opossum; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Chimeric - Didelphis virginiana.
OS   Chimeric - Canis sp.
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SEO1896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (RESI-) RESISTENTIA PHARM AB.

PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 2; 50pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2 and 4 of the opossum IgE and the heavy chain
CC   constant region 3 from the dog. It was shown to cause a stronger
CC   polyclonal anti-self IgE response than peptides consisting of the same
CC   regions from one mammal. Immunogenic peptides, particularly those
CC   consisting of different heavy chain constant regions, can be used for
CC   vaccination in humans, against bacterial and viral infections and
CC   allergies, such as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence 341 AA;

Qy   1 GMNLTWYRESK 11
         |||||||||||
Db 162 gmnltwyresk 172

Query Match   100.0%;   Score 63;   DB 21;   Length 341;
Best Local Similarity   100.0%;   Pred. No. 0.0035;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 7
AAW23067
ID   AAW23067 standard; Protein; 417 AA.
XX
AC   AAW23067;
XX
DT   19-FEB-1998 (first entry)
XX
DE   Canine IgE heavy chain constant region (exon 1-4 product).
XX
KW   IgE; immunoglobulin; antibody; heavy chain constant region;
KW   allergy; hypersensitivity; therapy; dog; antisense;
KW   immunomodulation.
XX
OS   Canis familiaris.
XX

Key   Location/Qualifiers
FH
FT   Misc-difference 55
FT   /note= "encoded by ACC"
FT   Misc-difference 56
FT   /note= "encoded by TAC"
FT   Misc-difference 67
FT   /note= "encoded by GCC"
FT   Misc-difference 83
FT   /note= "encoded by NNT"
FT   Misc-difference 174
FT   /note= "encoded by GGN"
FT   Misc-difference 175
FT   /note= "encoded by NNG"
FT   Misc-difference 176
FT   /note= "encoded by TGN"
FT   Misc-difference 203
FT   /note= "encoded by TCC"
FT   Misc-difference 204
FT   /note= "encoded by GAC"
XX
PN   WO9730156-A2.
XX
PD   21-AUG-1997.
XX
PF   14-FEB-1997;   97WO-US02322.
XX
PR   14-FEB-1996;   96US-0601197.

XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Harris RA, Mermer B, Siefring AE;
XX
DR   WPI; 1997-425031/39.
DR   N-PSDB; AAT79278.
XX
PT   Isolated canine IgE heavy chain constant region DNA - useful to
PT   develop products for treatment of canine allergies and for
PT   immunomodulation in dogs
XX
PS   Disclosure; Page 35-39; 59pp; English.
XX
CC   This polypeptide is encoded by exons 1-4 (see AAT79278) of canine
CC   IgE heavy chain constant region (epsilon) genomic DNA. Another
CC   polypeptide, comprising the exon 5 and 6 product, is given in
CC   AAW23068. Recombinant peptides encoded by exons 1-6 can be
CC   produced in eukaryotic or prokaryotic cells. Such peptides,
CC   and antibodies raised against them, are used in methods to treat
CC   the manifestation of allergy in dogs, e.g. to treatment Type I
CC   immediate hypersensitivity, and for immunomodulation.
XX
SQ   Sequence 417 AA;

Qy   1 GMNLTWYRESK 11
         |||||||||||
Db 240 gmnltwyresk 250

Query Match   100.0%;   Score 63;   DB 18;   Length 417;
Best Local Similarity   100.0%;   Pred. No. 0.0043;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 8
AAR97753
ID   AAR97753 standard; Protein; 426 AA.
XX
AC   AAR97753;
XX
DT   28-AUG-1996 (first entry)
XX
DE   Canine IgE.
XX
KW   IgE; immunoglobulin; antibody; vaccine; allergy; hypersensitivity.
XX
OS   Canis familiaris.
XX
PN   WO9614867-A1.
XX
PD   23-MAY-1996.
XX
PF   03-NOV-1995;   95WO-US13795.
XX
PR   09-NOV-1994;   94US-0336891.
PR   09-NOV-1994;   94US-0336583.
XX
PA   (MERI) MERCK & CO INC.
XX
PI   Hollis GF, Patel MD;
XX
DR   WPI; 1996-277321/28.
DR   N-PSDB; AAT29824.
XX
PT   New DNA encoding canine IgE and IgA - useful in vaccines, antisense
PT   therapy, assays, drug screening, etc.
XX
PS   Claim 11; Page 29-30; 49pp; English.
XX
CC   The canine IgE amino acid sequence (AAR97753) was deduced from
CC   an isolated gene (AAT29824) obtd. from a canine liver DNA library.
CC   The cloning of the IgE gene allows prodn. of large quantities of

[handwritten annotation in box: LXXYR.]

CC   reconbinant IgE using bacterial, yeast, mammalian, insect or
CC   viral systems. The IgE can be used in drug development (e.g.
CC   small molecule screening, assay development and anti-IgE
CC   antibody generation). Fragments of IgE can be used in vaccines
CC   or to prevent IgE-mediated hypersensitivity. The new sequence
CC   information permits targeted modulation of IgE-mediated immune
CC   reponses.
XX
SQ   Sequence 426 AA;

Qy   1 GMNLTWYRESK 11
            |||||||||||
Db   245 gmnltwyresk 255

Query Match   100.0%; Score 63; DB 17; Length 426;
Best Local Similarity 100.0%; Pred. No. 0.0044;
Matches 11; Conservative 11; Mismatches 0; Indels 0; Gaps 0;

RESULT 9
AAY69587
AAY69587 standard; peptide; 22 AA.
AC   AAY69587;
XX
DT   19-APR-2000 (first entry)
XX
DE   Peptide E3A.5-extended, SEQ ID NO:12.
XX
KW   Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW   allergy; canine; prophylaxis; treatment.
XX
OS   Canis familiaris.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107020.
PR   09-APR-1998; 98US-0058332.
PR   30-MAR-1999; 99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.
XX
New peptides, useful for generating canine autoantibodies against
immunoglobulin E (IgE) to treat allergic symptoms -

Example 1; Page 16; 38pp; English.

The invention relates to peptides (AAY69576-Y69586, AAY69598-Y69599) that
may be used to induce anti-IgE (immunoglobulin E) autoantibodies in a
dog. IgE is an important mediator of canine allergic responses, including
Type I immediate hypersensitivity. It is estimated that up to 30% of all
dogs suffer from allergies or allergy-related skin disorders, and
allergic dermatitis is thought to affect between 3 and 15% of dogs.
Common canine allergens include fleas, pollens, moulds and dust, with
flea allergy being the most frequent. Once IgE has been released in
response to allergen, the IgE binds to receptors on mast cells. The IgE
molecules can then become cross-linked, which induces the mast cells to
release histamine, which leads to the manifestation of allergic symptoms.
The peptides of the invention correspond to epitopes of canine IgE, or
conservative variations thereof, that are accessible in circulating IgE
but not in IgE bound to mast cells. The autoantibodies produced on
immunisation with such peptides therefore target only circulating IgE,
avoiding the possibility that they may cross-link bound IgE molecules,
which would exacerbate an ongoing allergic response even in the absence
of allergen. Anti-IgE autoantibody induction results in an immune

CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targeting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog. Sequences AAY69587-Y69588 represent peptides used
CC   in exemplifications of the present invention.
XX
SQ   Sequence 22 AA;

Qy   1 GMNLTWYRESK 11
            |||||||||||:
Db   1 gmnltwyrese 11

Query Match   93.7%; Score 59; DB 21; Length 22;
Best Local Similarity 90.9%; Pred. No. 0.00084;
Matches 10; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
AAY50882
ID   AAY50882 standard; peptide; 11 AA.
AC   AAY50882;
XX
DT   24-FEB-2000 (first entry)
XX
DE   Antibody 8H.8 binding substituted canine mimotope 1.
XX
KW   Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW   epitope; prophylaxis; treatment; mimotope.
XX
OS   Synthetic.
XX
PN   EP957111-A2.
XX
PD   17-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107035.
PR   09-APR-1998; 98US-0058331.
PR   30-MAR-1999; 99US-0281760.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-040833/04.
XX
PT   Binding proteins used for treatment or prophylaxis of canine allergy -

Example 2; Page 13; 30pp; English.

This invention describes a novel binding protein which specifically
binds to native canine free or B-cell bound IgE, and which doesn't bind
to IgE when the IgE is bound to mast cells. THe peptide products of the
invention have anti-IgE or anti-allergic activity. The antibodies bind to defined
epitopes on free or B-cell bound IgE molecules which have an important
role in allergic reaction. The specific binding proteins are used to
produce a pharmaceutical composition, preferably with a diluent, which
can be used for prophylaxis or treatment of canine allergy.
AAY50876-Y50900 represent peptide mimotopes used in the method of the
invention.
XX
SQ   Sequence 11 AA;

Query Match   85.7%; Score 54; DB 21; Length 11;
Best Local Similarity 90.9%; Pred. No. 0.0029;
Matches 10; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 GMNLTWYRESK 11
            ||||| ||||

Left column

```
Db   1 gmnltwsresk 11

RESULT 11
ID   AAY69592 standard; peptide; 12 AA.
XX
AC   AAY69592;
XX
DT   19-APR-2000 (first entry)
XX
DE   Canine substituted IgE-derived peptide, SEQ ID NO:17.
XX
KW   IgE; immunoglobulin E; allergy; canine; mutant; prophylaxis;
KW   treatment.
XX
OS   Canis familiaris.
XX
OS   Synthetic.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107020.
XX
PR   09-APR-1998; 98US-0058332.
PR   30-MAR-1999; 99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.
XX
PT   New peptides, useful for generating canine autoantibodies against
PT   immunoglobulin E (IgE) to treat allergic symptoms -
XX
PS   Example 6; Page 19; 38pp; English.
XX
CC   The invention relates to peptides (AAY69576-Y69586, AAY69598-Y69599) that
CC   may be used to induce anti-IgE (immunoglobulin E) autoantibodies in a
CC   dog. IgE is an important mediator of canine allergic responses, including
CC   Type I immediate hypersensitivity. It is estimated that up to 30% of all
CC   dogs suffer from allergies or allergy-related skin disorders, and
CC   allergic dermatitis is thought to affect between 3 and 15% of dogs.
CC   Common canine allergens include fleas, pollens, moulds and dust, with
CC   flea allergy being the most frequent. Once IgE has been released in
CC   response to allergen, the IgE binds to receptors on mast cells. The IgE
CC   molecules can then become cross-linked, which induces the mast cells to
CC   release histamine, which leads to the manifestation of allergic symptoms.
CC   The peptides of the invention correspond to epitopes of canine IgE, or
CC   conservative variations thereof, that are accessible in circulating IgE
CC   but not in IgE bound to mast cells. The autoantibodies produced on
CC   immunisation with such peptides therefore target only circulating IgE,
CC   avoiding the possibility that they may cross-link bound IgE molecules,
CC   which would exacerbate an ongoing allergic response even in the absence
CC   of allergen. Anti-IgE autoantibody induction results in an immune
CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targetting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog. Sequences AAY69590-Y69552 represent peptides
CC   compared with peptide AAY69584 in binding studies with the anti-canine
CC   IgE monoclonal antibody 8H.8 in an exemplification of the present
CC   invention.
XX
SQ   Sequence 12 AA;

Query Match      69.0%;  Score 43.5;  DB 21;  Length 12;
Best Local Similarity  83.3%;  Pred. No. 0.22;
Matches  10;  Conservative  0;  Mismatches  1;  Indels  1;  Gaps  1;
```

Right column

```
Qy   1 GMNLTW-YRESK 11
       |||||| ||||
Db   1 gmnltwseresk 12

RESULT 12
ID   AAR24251 standard; protein; 351 AA.
XX
AC   AAR24251;
XX
DT   23-NOV-1992 (first entry)
XX
DE   Vaccinia virus B18R protein.
XX
KW   recombinant; polyvalent vaccine; interleukin binding; IL-2; IL-6;
KW   multiple pathogens.
XX
OS   Human vaccinia virus.
XX
FH   Key             Location/Qualifiers
FT   Disulfide-bond  73..129
FT                   /note= "intradomain disulphide bridge"
FT   Disulfide-bond  172..221
FT                   /note= "intradomain disulphide bridge"
FT   Disulfide-bond  272..333
FT                   /note= "intradomain disulphide bridge"
FT   Region          117..119
FT                   /note= "N-linked carbohydrate addition site"
FT   Region          172..174
FT                   /note= "N-linked carbohydrate addition site"
FT   Region          251..253
FT                   /note= "N-linked carbohydrate addition site"
FT   Region          269..271
FT                   /note= "N-linked carbohydrate addition site"
FT   Region          321..323
FT                   /note= "N-linked carbohydrate addition site"
XX
PN   WO9207944-A.
XX
PD   14-MAY-1992.
XX
PF   28-OCT-1991; 91WO-GB01882.
XX
PR   26-OCT-1990; 90GB-0023352.
XX
PA   (LYNX-) LYNXVALE LTD.
XX
PI   Smith GL;
XX
DR   WPI; 1992-183679/22.
DR   N-PSDB; AAQ25097.
XX
PT   New vaccinia vector genes and expression products - useful in
PT   diagnosis and treatment of inflammatory conditions
XX
PS   Disclosure; Fig 1C; 73pp; English.
XX
CC   The sequence is that of the vaccinia virus B18R protein
CC   which is involved in the binding of interleukin (IL)-2 and/or IL-6
CC   and in preventing them from reaching their natural receptors, hence
CC   the inflammatory response is diminished and viral replication is
CC   enhanced. Its gene sequence may be inactivated or all or part of it
CC   deleted to allow greater attenuation of the virus, enhanced
CC   immunogenicity of the recombinant vaccinia virus and/or provide
CC   sites for the insertion of heterologous genes. This recombinant
CC   vaccinia virus may be used as a live vaccine in both human and
CC   veterinary medicine. See also AAR24250 and AAR24252.
XX
SQ   Sequence 351 AA;

Query Match      65.1%;  Score 41;  DB 13;  Length 351;
```

Best Local Similarity   55.6%;   Pred. No. 24;
Matches   5;   Conservative   4;   Mismatches   0;   Indels   0;   Gaps   0;
XX

Qy   3 NLTWYRESK 11
          |:||||::|
Db   182 nitwykdnk 190

RESULT 13
AAR41542
ID   AAR41542 standard; Protein; 351 AA.
XX
AC   AAR41542;
XX
DT   25-MAR-1994 (first entry)
XX
DE   B15R product.
XX
KW   B15R; interleukin-1 beta; IL-1j; diagnosis; kit; inflammation;
KW   rheumatoid arthritis; septic shock; cancer; fever; bowel disease;
KW   graft versus host disease; diabetes; expression; B18R.
XX

OS   Vaccinia virus.

FH   Key   Location/Qualifiers
FH
FT   Peptide   9..19
FT   /label= possible_sig_peptide
FT   Modified-site   117..119
FT   /label= possible_N-linked_carbohydrate_site
FT   Modified-site   184..186
FT   /label= possible_N-linked_carbohydrate_site
FT   Modified-site   261..263
FT   /label= possible_N-linked_carbohydrate_site
FT   Modified-site   269..271
FT   /label= possible_N-linked_carbohydrate_site
FT   Modified-site   321..323
FT   /label= possible_N-linked_carbohydrate_site
FT   Disulfide-bond 73
FT   /label= possible_linking_site
FT   Disulfide-bond 129
FT   /label= possible_linking_site
FT   Disulfide-bond 172
FT   /label= possible_linking_site
FT   Modified-site   221
FT   /label= possible_linking_site
FT   Disulfide-bond 272
FT   /label= possible_linking_site
FT   Disulfide-bond 333
FT   /label= possible_linking_site
XX

PR   WO9318153-A.
XX
PD   16-SEP-1993.
XX
PF   05-MAR-1993; 93WO-GB00460.
XX
PR   05-MAR-1992; 92GB-0004780.
XX
PA   (SMIT/) SMITH G L.
PI   Smith GL;
XX
DR   WPI; 1993-303467/38.
DR   N-PSDB; AAQ46244.
XX
PT   New vaccinia virus B15R nucleotide sequence - used to prepare
PT   prods. to treat conditions with interleukin-I beta mediated
PT   symptoms
XX
PS   Disclosure; Fig 1C; 91pp; English.
XX
CC   The sequence (AAQ46243) encodes a protein that specifically binds to
CC   IL-1-beta and can be used to block the activity of IL-1-beta and treat

CC   IL-1-beta mediated conditions. The sequence can also be used as part
CC   of a kit to detect and mesure IL-1B. The sequence (AAQ46244) encodes a
CC   similar protein and recombinant vvs in which expression of B15R and
CC   B18R is inactivated. To introduce restriction sites close to the
CC   B18R is inactivated. To introduce restriction sites close to the
CC   initiator methionine of B15R, a PCR copy of the 5' region of the
CC   ORF was constructed using primers (AAQ46241-42).
XX
SQ   Sequence   351 AA;

Query Match   65.1%;   Score 41;   DB 14;   Length 351;
Best Local Similarity   55.6%;   Pred. No. 24;
Matches   5;   Conservative   4;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   3 NLTWYRESK 11
          |:||||::|
Db   182 nitwykdnk 190

RESULT 14
AAY93969
ID   AAY93969 standard; Protein; 574 AA.
XX
AC   AAY93969;
XX
DT   03-OCT-2000 (first entry)
XX
DE   Amino acid sequence of a partial rat antigen Ag2117.
XX
KW   Antigen; Ag2117; pancreatic cell; monoclonal antibody; surface antigen.
XX
OS   Rattus sp.
XX
PN   WO200037503-A1.
XX
PD   29-JUN-2000.
XX
PF   22-DEC-1999;   99WO-US30741.
XX
PR   22-DEC-1998;   98US-0218539.
XX
PA   (RAVE-) RAVEN BIOTECHNOLOGIES INC.
XX
PI   Mather JP, Bald LN, Roberts PR, Stephan JF;
XX
DR   WPI; 2000-452179/39.
DR   N-PSDB; AAA57480.
XX
PT   Immunizing a host mammal to produce population of monoclonal antibodies
PT   that bind to antigens of specific cell type comprises introducing
PT   viable cells with serum-free surfaces ...
XX
PS   Example 10; Fig 10; 61pp; English.
XX
CC   The present sequence represents rat antigen Ag2117. The antigen is
CC   representative of pancreatic cell lines. The antibody that recognises
CC   Ag2117 was identified using the method of the invention. The
CC   specification describes a method for immunizing a host mammal to produce
CC   a population of monoclonal antibodies that bind to antigens
CC   representative of a specific cell type that are heterologous to the
CC   host mammal. The method comprises introducing into the mammal intact
CC   and viable cells with surfaces which are free of serum. The methods
CC   are useful for producing populations of antibodies that bind to antigens
CC   representative of a specific cell type, specifically antibodies that
CC   are tissue-selective, sub-tissue selective or cell-type specific. These
CC   antibodies facilitate the identification of novel antigens and the
CC   delineation of the combination of surface antigens on a specific cells
CC   type. The antibodies or hybridomas produced using the methods have
CC   diagnostic and therapeutic applications.
XX
SQ   Sequence   574 AA;

```
Query Match      65.1%;     Score 41;  DB 21;  Length 574;
Best Local Similarity   66.7%;  Pred. No. 41;
Matches   6; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy      3 NLTWYRESK 11
          |:||||| |
Db    167 nitwyrngk 175
```

RESULT 15
XX
ID   AAB76858 standard; Protein; 582 AA.
XX
AC   AAB76858;
XX
DT   12-APR-2001 (first entry)
XX
DE   Human lung tumour protein related protein sequence SEQ ID NO:334.
XX
KW   Human; lung cancer; lung tumour protein; gene therapy;
KW   lung cancer antigen; lung tumour-specific antigen; diagnosis; vaccine;
KW   cytostatic; antisense inhibition.
XX
OS   Homo sapiens.
XX
PN   WO200100828-A2.
XX
PD   04-JAN-2001.
XX
PF   30-JUN-2000; 2000WO-US18061.
XX
PR   30-JUN-1999; 99US-0346492.
PR   15-OCT-1999; 99US-0419356.
PR   17-DEC-1999; 99US-0468867.
PR   30-DEC-1999; 99US-0476300.
PR   06-MAR-2000; 2000US-0519642.
PR   22-MAR-2000; 2000US-0533077.
PR   10-APR-2000; 2000US-0545259.
PR   27-APR-2000; 2000US-0560406.
PR   05-JUN-2000; 2000US-0589184.
XX
PA   (CORI-) CORIXA CORP.
XX
PI   Wang T, Bangur CS, Lodes MJ, Fanger GR, Vedvick TS, Carter D;
PI   Retter MW, Manlon J;
XX
DR   WPI; 2001-071488/08.
XX
PT   lung tumour-associated proteins and the nucleic acids that encode them,
PT   useful for preventing, diagnosing and treating lung cancer -
XX
PS   Example 1; Page 262-264; 436pp; English.
XX
CC   The present invention describes immunogenic portions of lung tumour-
CC   associated proteins (I) and the nucleic acids (NAs) that encode them.
CC   (I) have cytostatic activity and can be used in gene therapy, antisense
CC   inhibition and in vaccines. The NAs and the lung tumour-associated
CC   proteins they encode may be used in the prevention, treatment and
CC   diagnosis of diseases associated with their inappropriate expression,
CC   especially lung cancer. For example, the NAs may be administered to
CC   treat diseases by rectifying mutations or deletions in a patient's genome
CC   that affect the activity of the protein by expressing inactive proteins
CC   or to supplement the patients own production of (I). Additionally, the
CC   NAs may be used to produce the lung-tumour associated protein, according
CC   to standard recombinant DNA methodology. Conversely, antisense NA
CC   molecules may be administered to down regulate protein expression by
CC   binding with the cells own genes and preventing their expression. The NA
CC   and complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar NA sequences in
CC   samples, and hence which patients may be in need of treatment for lung
CC   cancer. The (I) may be used as antigens in the production of antibodies
CC   and in assays to identify modulators (agonists and antagonists) of the
CC   expression and activity of the protein. AAF68083 to AAF68878 and
CC   AAB76848 to AAB76878 represent human lung tumour protein related
CC   nucleotide and protein sequences which are used in the exemplification
CC   of the present invention.
XX
SQ   Sequence   582 AA;

```
Query Match      65.1%;     Score 41;  DB 22;  Length 582;
Best Local Similarity   66.7%;  Pred. No. 42;
Matches   6; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy      3 NLTWYRESK 11
          |:||||| |
Db    166 nitwyrngk 174
```

RESULT 16
XX
ID   AAW06891 standard; Protein; 583 AA.
XX
AC   AAW06891;
XX
DT   18-MAR-1997 (first entry)
XX
DE   Human activated leukocyte-cell adhesion molecule ALCAM.
XX
KW   Activated luekocyte-cell adhesion molecule; ALCAM; CD6 ligand;
KW   thymic epithelial cell; inflammation; allograft rejection;
KW   neurodegenerative disease.
XX
OS   Homo sapiens.
XX
FH   Key   Location/Qualifiers
FH
FT   Peptide   1..27
FT   /label= Sig_peptide
FT   /note= "putative N-glycosylation site"
FT   Protein   28..583
FT   /label= Mat_protein
FT   Domain   28..527
FT   /label= Extracellular_domain
FT   /note= "the extracellular doamin can be subdivid"
FT   Domain   528..551
FT   /label= Transmembrane_domain
FT   Domain   552..583
FT   /label= Extracellular_domain
FT   Modified-site   91
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Modified-site   95
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Misc-difference 167
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Modified-site   258
FT   /label= Asn, Ser
FT   /note= "variation owing to polymorphism in HL60
FT   and T-cell derived clones"
FT   Modified-site   265
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Misc-difference 301
FT   /label= Met, Thr
FT   /note= "variation owing to polymorphism in HL60
FT   and T-cell derived clones"
FT   Modified-site   306
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Modified-site   337
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"
FT   Modified-site   361
FT   /label= Glycosylation
FT   /note= "putative N-glycosylation site"

```
FT   Modified-site
FT      /label= Glycosylation
FT      /note= "putative N-glycosylation site"
FT   Modified-site   457
FT      /label= Glycosylation
FT    480
FT      /note= "putative N-glycosylation site"
FT   Modified-site   499
FT      /label= Glycosylation
FT      /note= "putative N-gklycosylation site"
XX
PN   WO9634880-A1.
XX
PD   07-NOV-1996.
XX
PF   29-APR-1996;  96WO-US06010.
XX
PR   01-MAY-1995;  95US-0432016.
XX
DR   WPI; 1996-506097/50.
DR   N-PSDB; AAT46075.
XX
PA   (UVDU-) UNIV DUKE.
PA   (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
PI   Aruffo A, Bowen MA, Haynes BF, Marquardt H, Patel D;
XX
PT   CD6 ligands present on surface of thymic epithelial cells - used to
PT   develop prods. for treating e.g. inflammation, organ allograft
PT   rejection or neuro:degenerative diseases
XX
PS   Example 4; Fig 29; 112pp; English.
XX
CC   Human activated leukocyte cell adhesion molecule (AAW06891) or ALCAM
CC   is a CD6 ligand present on the surface of thymic epithelial cells,
CC   monocytes, activated T-cells, and other cells. Its amino acid
CC   sequence was deduced from cDNA clones (AAT46075) obtd. from HL60
CC   and T-cell lines. ALCAM polypeptides can be produced, e.g. as a
CC   fusion proteins, in transformed host cells. They are useful for
CC   inhibiting CD6/CD6 interactions between cells and to screen for
CC   agents that modulate this interaction. Inhibitors can be used to
CC   treat e.g. inflammation, multiple sclerosis, inflammatory uveitis,
CC   rheumatoid arthritis, T-cell mediated vasculitis syndromes, organ
CC   allograft rejection and neurodegenerative diseases.
XX
SQ   Sequence   583 AA;

Query Match         65.1%;  Score 41;  DB 17;  Length 583;
Best Local Similarity   66.7%;  Pred. No. 42;
Matches    6;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

        3 NLTWYRESK 11
          |:||| |
      167 nltwyrngk 175

RESULT 17
ID   AAR97230 standard; Protein; 583 AA.
XX
AC   AAR97230;
XX
DT   08-OCT-1996  (first entry)
XX
DE   Stem cell marker HCAPro.1.
XX
KW   Stem cell marker; haematopoietic cell antigen; HCA; HCASeq.1;
XX
OS   Homo sapiens.
XX
PN   EP716146-A2.
XX
```

```
PD   12-JUN-1996.
XX
PF   02-FEB-1995;  95EP-0300661.
XX
PR   06-DEC-1994;  94US-0352323.
XX
PA   (SYST-) SYSTEMIX INC.
XX
PI   Gearing D, Uchida N, Yang Z;
XX
DR   WPI; 1996-269990/28.
DR   N-PSDB; AAT28819.
XX
PS   Claim 8; Fig 4A; 17pp; English.
XX
CC   HCAPro.1 (AAR97230) and HCAPro.2 (AAR97231) are novel human
CC   haematopoietic cell antigens associated with stem cells. They
CC   are the respective products of nucleic acid sequences HCASeq.1
CC   (AAT28819) and HCASeq.2 (AAT28820) obtd. from human CD34+ bone marrow.
CC   Expression systems are provided for prodn. of recombinant HCA
CC   proteins. The proteins are useful as stem cell marker proteins in
CC   functional studies, and can also be used to produce antibodies
CC   that allow the purification of stem cells from haematopoietic and
CC   other sources.
XX
PT   DNA encoding human haematopoietic cell antigen proteins - useful as
PT   stem cell marker proteins in functional studies and for antibody
PT   prodn.
XX
SQ   Sequence   583 AA;

Query Match         65.1%;  Score 41;  DB 17;  Length 583;
Best Local Similarity   66.7%;  Pred No. 42;
Matches    6;  Conservative   1;  Mismatches   2;  Indels   0;  Gaps   0;

Qy       3 NLTWYRESK 11
           |:||| |
Db     167 nltwyrngk 175

RESULT 18
ID   AAW47088 standard; Protein; 583 AA.
XX
AC   AAW47088;
XX
DT   20-JUL-1998  (first entry)
XX
DE   Activated leukocyte-cell adhesion molecule (ALCAM).
XX
KW   ALCAM; activated leukocyte-cell adhesion molecule; CD6 ligand;
KW   antibody; human.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..27
FT      /label= Sig_peptide
FT   Protein         28..583
FT      /label= Mat_protein
FT   Domain          28..527
FT      /note= "extracellular domain"
FT   Domain          528..551
FT      /note= "transmembrane domain"
FT   Domain          552..583
FT      /note= "cytoplasmic domain"
FT   Modified-site   91
FT      /note= "N-glycosylated"
FT   Modified-site   95
FT      /note= "N-glycosylated"
FT   Modified-site   167
FT      /note= "N-glycosylation"
```

FT   Misc-difference   258
FT   /label= Asn, Ser
FT   /note= "encoded by ART (polymorphism in nucleotide
FT   sequence"
FT   Modified-site   265
FT   /note= "N-glycosylated"
FT   Misc-difference   301
FT   /label= Met, Thr
FT   /note= "encoded by AYG (polymorphism in nucleotide
FT   sequence)"
FT   Modified-site   306
FT   /note= "encoded by CAR"
FT   Modified-site   337
FT   /note= "N-glycosylated"
FT   Modified-site   361
FT   /note= "N-glycosylated"
FT   Misc-difference   447
FT   /note= "encoded by CAR"
FT   Modified-site   457
FT   /note= "N-glycosylated"
FT   Modified-site   480
FT   /note= "N-glycosylated"
FT   Modified-site   499
FT   /note= "N-glycosylated"
XX
PN   WO9803551-A1.
XX
PD   29-JAN-1998.
XX
PF   17-JUL-1997; 97WO-US12511.
XX
PR   18-JUL-1996; 96US-0684594.
XX
PA   (BRIM) BRISTOL-MYERS SQUIBB CO.
PA   (UYDU-) UNIV DUKE.
XX
PI   Aruffo A, Bowen MA, Haynes BF, Marquardt H, Patel D;
PI   Sladak AW;
XX
DR   WPI; 1998-120703/11.
DR   N-PSDB; AAV13954.
XX
PT   Activated leukocyte-cell adhesion molecule (ALCAM), a new CD6 ligand
PT   - useful for screening compounds which inhibit interaction of ALCAM
PT   and CD6 ligand antibodies
XX
PS   Claim 3: Fig 29; 130pp; English.
XX
CC   This polypeptide is a human CD6 ligand (see AAW47088), designated
CC   activated leukocyte-cell adhesion molecule (ALCAM), that is present
CC   on the surface of thymic epithelial cells, monocytes, activated
CC   T cells and a variety of other cells. Its amino acid sequence was
CC   deduced from cDNA clones (see AAV13954) isolated from human T cell
CC   and HL60 cDNA libraries. It is the human homologue of chicken
CC   BEN, and shows homology to neurolin, RAGE and MUC18. Anti-CD6
CC   ligand antibodies, or their binding fragments, are useful for
CC   inhibiting binding of CD6 present on the surface of a first cell
CC   to that present on a second cell. CD6 ligands and anti-CD6 ligand
CC   antibodies are also useful in screening test compounds for the
CC   ability to inhibit binding of CD6 ligand to an anti-CD6 ligand
CC   antibody (all claimed).
XX
SQ   Sequence   583 AA;

Qy   3 NLTWYRESK 11
            |:||| |
Db   167 nltwyrngk 175

Query Match   65.1%; Score 41; DB 19; Length 583;
Best Local Similarity 66.7%; Pred. No. 42;
Matches   6; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

RESULT 19
AAY69591
ID   AAY69591 standard; peptide; 11 AA.
XX
AC   AAY69591;
XX
DT   19-APR-2000   (first entry)
XX
DE   Feline IgE-derived peptide, SEQ ID NO:16.
XX
KW   IgE; immunoglobulin E; feline; allergy; canine; prophylaxis; treatment.
XX
OS   Felis catus.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107020.
XX
PR   09-APR-1998; 98US-0058332.
PR   30-MAR-1999; 99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.
XX
PT   New peptides, useful for generating canine autoantibodies against
PT   immunoglobulin E (IgE) to treat allergic symptoms -
XX
PS   Example 6; Page 19; 38pp; English.
XX
CC   The invention relates to peptides (AAY69576-Y69586, AAY69598-Y69599) that
CC   may be used to induce anti-IgE (immunoglobulin E) autoantibodies in a
CC   dog. IgE is an important mediator of canine allergic responses, including
CC   Type I immediate hypersensitivity. It is estimated that up to 30% of all
CC   dogs suffer from allergies or allergy-related skin disorders, and
CC   allergic dermatitis is thought to affect between 3 and 15% of dogs.
CC   Common canine allergens include fleas, pollens, moulds and dust, with
CC   flea allergy being the most frequent. Once IgE has been released in
CC   response to allergen, the IgE binds to receptors on mast cells. The IgE
CC   molecules can then become cross-linked, which induces the mast cells to
CC   release histamine, which leads to the manifestation of allergic symptoms.
CC   The peptides of the invention correspond to epitopes of canine IgE, or
CC   conservative variations thereof, that are accessible in circulating IgE
CC   but not in IgE bound to mast cells. The autoantibodies produced on
CC   immunisation with such peptides therefore target only circulating IgE,
CC   avoiding the possibility that they may cross-link bound IgE molecules,
CC   which would exacerbate an ongoing allergic response even in the absence
CC   of allergen. Anti-IgE autoantibody induction results in an immune
CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targeting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog. Sequences AAY69590-Y69592 represent peptides
CC   compared with peptide AAY69584 in binding studies with the anti-canine
CC   IgE monoclonal antibody 8H.8 in an exemplification of the present
CC   invention.
XX
SQ   Sequence   11 AA;

Qy   1 GMNLTWYRES 10
            || ||| ||:
Db   1 gmtltwsren 10

Query Match   63.5%; Score 40; DB 21; Length 11;
Best Local Similarity 70.0%; Pred. No. 0.8;
Matches   7; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

RESULT 20
XX
ID   AAY50881 standard; peptide; 11 AA.
XX
AC   AAY50881;
XX
DT   24-FEB-2000 (first entry)
XX
DE   Antibody 8H.8 binding feline mimotope 1.
XX
KW   Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW   epitope; prophylaxis; treatment; mimotope.
XX
OS   Synthetic.
XX
PN   EP957111-A2.
XX
PD   17-NOV-1999.
XX
PF   09-APR-1999;   99EP-0107035.
XX
PR   30-MAR-1999;   99US-0281760.
     09-APR-1998;   98US-0058331.
XX
DR   WPI: 2000-040833/04.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PT   Binding proteins used for treatment or prophylaxis of canine allergy -
     Example 2; Page 13; 30pp; English.
XX
CC   This invention describes a novel binding protein which specifically
CC   binds to native canine free or B-cell bound IgE, and which doesn't bind
CC   to IgE when the IgE is bound to mast cells. The peptide products of the
CC   invention have anti-allergic activity. The antibodies bind to defined
CC   epitopes on free or B-cell bound IgE molecules which have an important
CC   role in allergic reaction. The specific binding proteins are used to
CC   produce a pharmaceutical composition, preferably with a diluent, which
CC   can be used for prophylaxis or treatment of canine allergy.
CC   AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC   invention.
XX
SQ   Sequence   11 AA;

Query Match   63.5%; Score 40; DB 21; Length 11;
Best Local Similarity   70.0%; Pred. No. 0.8;
Matches   7; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 GMNLTWYRES 10
          ||=||=||:|
Db   1 gmtltwsren 10

RESULT 21
XX
ID   AAB06204 standard; protein; 343 AA.
XX
AC   AAB06204;
XX
DT   22-NOV-2000 (first entry)
XX
DE   Platypus IgE heavy chain constant regions 2, 3 and 4.
XX
KW   Platypus; immunoglobulin E; IgE; vaccination; infection; allergy;
KW   asthma; eczema; immunogenic peptide.
XX
OS   Ornithorhynchus anatinus.
XX

FH   Key   Location/Qualifiers
FT   Misc-difference 1..343
FT   /label= OTHER
FT   /note= "Xaa=unknown"
XX
PN   WO200025722-A2.
XX
PD   11-MAY-2000.
XX
PF   21-OCT-1999;   99WO-SE01896.
XX
PR   02-NOV-1998;   98US-0106652.
PR   22-SEP-1999;   99US-0401636.
XX
PA   (REST-) RESISTENTIA PHARM AB.
XX
PI   Hellman LT;
XX
DR   WPI; 2000-365342/31.
XX
PT   Immunogenic polypeptides useful for preventing the harmful effects of
PT   immunoglobulin E in mammals -
XX
PS   Disclosure; Fig 2; 5pp; English.
XX
CC   The present sequence is an immunogenic peptide consisting of the heavy
CC   chain constant regions 2, 3 and 4 of the platypus IgE. It was used to
CC   construct a number of immunogenic peptides which consisted of regions of
CC   IgE from different mammals, which appear to cause a stronger polyclonal
CC   anti-self IgE response than peptides consisting of the same regions from
CC   one mammal. Immunogenic peptides, particularly those consisting of
CC   different heavy chain constant regions, can be used for vaccination
CC   in humans, against bacterial and viral infections and allergies, such
CC   as asthma, fur, pollen and food allergies and eczema.
XX
SQ   Sequence   343 AA;

Query Match   63.5%; Score 40; DB 21; Length 343;
Best Local Similarity   70.0%; Pred. No. 35;
Matches   7; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 GMNLTWYRES 10
          ||:|||||||
Db   163 gmevtwsres 172

RESULT 22
XX
ID   AAG93603 standard; protein; 107 AA.
XX
AC   AAG93603;
XX
DT   14-SEP-2001 (first entry)
XX
DE   Human anti-Rh(D) chain I16 protein sequence.
XX
KW   Human; anti-Rh(D) antibody; Rh(D) binding protein; immunostimulant;
KW   red blood cell; Rh phenotype; diagnosis; therapeutic.
XX
OS   Homo sapiens.
XX
PN   US625455-B1.
XX
PD   03-JUL-2001.
XX
PF   29-JAN-1999;   99US-0240274.
XX
PR   11-OCT-1996;   96US-0028550.
PR   10-APR-1998;   98US-0081380.
PR   27-JUN-1997;   97US-0884045.
XX
PA   (UYPE-) UNIV PENNSYLVANIA.

XX
PI Siegel DL;
XX
DR WPI; 2001-388931/41.
DR N-PSDB; AAH68660.
XX
PT New isolated protein, preferably a human anti-Rh(D) antibody for use in
PT diagnostics requiring a human instead of an animal antibody and in
PT therapeutic medicine -
XX
PS Claim 1; Column 44; 162pp; English.
XX
CC The present invention describes an isolated Rh(D) binding protein,
CC preferably a human antibody, (I) having an amino acid sequence comprising
CC one of the sequences (S) given in AAG93558 to AAG93669. (I) has
CC immunostimulant activity, and can be used as an immune system stimulant.
CC (I) can be used in diagnostic and therapeutic medicine. The antibodies
CC are used in diagnostics that require human antibodies instead of animal
CC antibodies, such as determine the Rh phenotype of human red blood cells.
CC AAH68615 to AAH68726 represent the nucleotide sequence which encode
CC AAG93558 to AAG93669. AAG93670 to AAG93697 represent anti-Rh(D) heavy
CC chain CDR3 amino acid sequences which are given in the exemplification
CC of the present invention.
XX
SQ Sequence   107 AA;

Query Match   60.3%; Score 38; DB 22; Length 107;
Best Local Similarity   50.0%; Pred. No. 22;
Matches   5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy   1 GMNLTWYRES 10
             | ||.|:::
Db   29 gfnlnwyqt 38

RESULT 23
AAW97813
ID AAW97813 standard; protein; 130 AA.
XX
AC AAW97813;
XX
DT 21-MAY-1999 (first entry)
XX
DE Bovine butyrophilin IgI domain.
XX
KW Butyrophilin; cattle; bovine; multiple sclerosis; diagnosis;
KW vaccine; transgenic animal; food; dairy product; milk.
XX
OS Bos taurus.
XX
CC WO9905162-A1.
XX
PN 04-FEB-1999.
XX
PD 22-JUL-1998; 98WO-US15062.
XX
PF 22-JUL-1997; 97US-0053376.
XX
PR
PA (UYMA-) UNIV MARYLAND BALTIMORE.
XX
PI Linington C, Mather IH;
XX
DR WPI; 1999-142845/12.
XX
PT New modified butyrophilin - not cross-reactive with myelin
PT oligo-dendrocyte glycoprotein, useful in dairy products, vaccines,
PT and assays for susceptibility to multiple sclerosis
XX
PS Disclosure; Page 38-39; 53pp; English.
XX
CC This polypeptide comprises the intermediate-type immunoglobulin
CC like domain (IgI of bovine butyrophilin (BTN, see AAW97812), a major

CC component of the milk fat globule membrane and member of the
CC immunoglobulin (Ig) superfamily. The IgI domain of BTN has a high
CC degree of identity with the IgV domain of myelin/oligodendrocyte
CC glycoprotein (MOG, see AAW97817). Immunisation with exogenous BTN
CC induces the production of antibodies which cross-react with
CC endogenous MOG, and this predisposes animals to pathological
CC autoimmune responses such as multiple sclerosis (MS). The
CC invention is directed to methods of modifying the BTN protein to
CC eliminate sequences that elicit this immune response to MOG, e.g.
CC by removing or modifying the IgI domain of a non-human BTN (nhBTN)
CC in a diary product, or by producing a humanised chimeric BTN
CC composed of an exoplasmic domain comprising a human IgI domain and
CC nhBTN (preferably bovine) IgC1 domain, and a cytoplasmic domain
CC comprising a nhBTN (preferably bovine) B30.2 region. The invention
CC includes genetic manipulation of the endogenous BTN gene of an
CC animal such that the human IgI domain is introduced into the
CC germline of the animal, and the animal expresses the humanised
CC protein in its milk. Dairy products including modified nhBTN or
CC chimeric BTN are also provided. A diagnostic method for detecting
CC susceptibility to MS in a patient from injecting bovine BTN is
CC claimed. Identification of a human patient at risk of developing a
CC pathogenic autoimmune response to MOG on exposure to nhBTN is also
CC possible, by screening the patient to identify polymorphisms in the
CC gene with domains which are at least 50% identical to nhBTN IgI
CC domains (claimed). Peptides derived from nhBTN, and constructs
CC based on nucleic acids encoding nhBTN, are used as vaccines to
CC induce Immunological tolerance to nhBTN in humans (claimed).
XX
SQ Sequence   130 AA;

Query Match   60.3%; Score 38; DB 20; Length 130;
Best Local Similarity   66.7%; Pred. No. 27;
Matches   6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   1 GMNLTWYRE 9
             | ||.|:.||
Db   34 gmelrwfre 42

RESULT 24
AAB60775
ID AAB60775 standard; peptide; 193 AA.
XX
AC AAB60775;
XX
DT 27-MAR-2001 (first entry)
XX
DE Gene 47 related peptide #1.
XX
KW Secreted protein; gene therapy; vaccine; cancer; leukemia;
KW autoimmune disease; allergy; inflammation; graft rejection;
KW hyperproliferation; cardiovascular; infection.
XX
OS Homo sapiens.
XX
CC WO20007653I-A1.
XX
PN 21-DEC-2000.
XX
PD 01-JUN-2000; 2000WO-US15137.
XX
PF 11-JUN-1999; 99US-0138625.
XX
PR
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM, Komatsoulis GA;
XX
DR WPI; 2001-071148/08.
XX
PT Nucleic acids encoding 47 human secreted polypeptides, useful for
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's
PT disease and diabetic retinopathy -

```
XX
PS   Disclosure; Page 523-524; 525pp; English.
XX
CC   The present invention relates to 26 secreted human proteins. The
CC   proteins may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate polypeptide expression.
CC   For example, they may be used in gene therapy or in vaccines.
CC   Typical of diseases which are potentially treatable are cancers
CC   (including leukemia), autoimmune diseases, allergies, inflammation,
CC   graft rejection, hyperproliferation, cardiovascular diseases
CC   (particularly critical limb ischemia and coronary disease) and any
CC   involving abnormal angiogenesis, neurodegeneration and/or
CC   infectious diseases.
XX
SQ   Sequence    193 AA;

Query Match     60.3%;  Score 38;  DB 22;  Length 193;
Best Local Similarity   66.7%;               Pred. No. 42;
Matches    6;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

         1 GMNLTWYRE 9
           ||  | |:||
        19 gmelrwfre 27

RESULT  25
ID   AAW10505 standard; Protein; 196 AA.
XX
AC   AAW10505;
XX
DT   12-SEP-1997  (first entry)
XX
DE   Alphalalpha2 region of Class II NOD mouse MHC (IAg7).
XX
KW   Soluble; fusion; major histocompatibility complex; MHC; region;
KW   heterodimer; complex; alpha2; antigen; binding groove; tolerance;
KW   autoantigen; disease; insulin dependent; diabetes mellitus; IDDM;
KW   antagonist; T cell; anergy; presenting cell; NOD mouse;
KW   Class II; alpha.
XX
OS   Mus spp.
XX
XX
PN   WO9640944-A2.
XX
PD   19-DEC-1996.
XX
PF   07-JUN-1996;  96WO-US10102.
XX
PR   27-OCT-1995;  95US-0005964.
PR   07-JUN-1995;  95US-0480002.
PR   07-JUN-1995;  95US-0482133.
PR   07-JUN-1995;  95US-0483241.
XX
PA   (ANER-) ANERGEN INC.
PA   (ZYMO ) ZYMOGENETICS INC.
XX
PI   Deshpande S,  Gross JA,  Kindsvogel W,  Reich EP,  Sheppard PO;
XX
DR   WPI; 1997-052337/05.
DR   N-PSDB; AAT60698.
XX
PT   Novel fused major histocompatibility complex:antigenic peptide
PT   complex - useful to induce tolerance to an autoantigen-related
PT   disease e.g. insulin-dependent diabetes mellitus
XX
PS   Example 3; Pages 132-133; 142pp; English.
XX
CC   A novel soluble fused major histocompatibility complex (MHC)
CC   heterodimer:peptide complex, comprises DNA encoding 1st and 2nd
CC   MHC domains, e.g. the present sequence, linked by DNA encoding a
CC   5-25 residue linker, and a DNA encoding an antigenic peptide able
```

```
CC   to associate with a peptide binding groove of the MHC molecule,
CC   linked in frame to the DNA encoding the 2nd domain by a DNA
CC   encoding a 5-25 residue linker. The complex can be used to induce
CC   immunological tolerance in adults susceptible to, or suffering from
CC   an autoantigen related disease, e.g. insulin dependent diabetes
CC   mellitus (IDDM), by antagonising the binding of particular T cells
CC   and antigen presenting cells, to induce anergy (immunological
CC   non-responsiveness) in the targeted T cell. As the heterodimers and
CC   corresponding antigen are permanently linked into a single chain,
CC   obviating the requirement for complex heterodimer truncation or
CC   formation, the complex eliminates inefficient and non-specific
CC   peptide loading.
XX
SQ   Sequence    196 AA;

Query Match     60.3%;  Score 38;  DB 18;  Length 196;
Best Local Similarity   60.0%;               Pred. No. 42;
Matches    6;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

Qy    2 MNLTWYRESK 11
        |:|| | ||
Db  121 lnitwlrnsk 130

Search completed: April 21, 2002, 10:14:14
Job time: 184 sec
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:46:37 ; Search time 26.44 Seconds
                                             (without alignments)
                                             19.611 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 52463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 47201

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   A_Geneseq_1101.*
                 1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                 2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                 3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                 4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                 5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                 6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                 7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                 8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                 9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Minimum DB seq length: 0
Maximum DB seq length: 7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	35	100.0	7	21	AAY69584	Immunogenic peptid
2	35	100.0	7	21	AAY50878	Antibody 8H.8 bind
3	23	65.7	7	22	AAU03292	Fruit fly G protei
4	22	62.9	7	20	AAY48904	Membrane dipeptida
5	22	62.9	7	21	AAB21922	Human prostate-hom
6	22	62.9	7	22	AAE06498	Mouse prostate hom
7	21	60.0	6	21	AAB03076	Oyster-derived HIV
8	21	60.0	4	21	AAB03078	Peptide SEQ ID NO:
9	20	57.1	6	21	AAB03077	HIV-1 protease inh
10	20	57.1	6	21	AAB03075	HIV-1 protease inh
11	19	54.3	5	18	AAW12480	Interleukin-6 anta

RESULT   1
AAY69584

ID   AAY69584 standard; peptide; 7 AA.
XX
AC   AAY69584;
XX
DT   19-APR-2000 (first entry)
XX
DE   Immunogenic peptide SEQ ID NO:9, for anti-IgE autoantibody generation.
XX
KW   Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW   allergy; canine; prophylaxis; treatment.
XX
OS   Synthetic.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999;   99EP-0107020.
XX
PR   09-APR-1998;   98US-0058332.
PR   30-MAR-1999;   99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.

ALIGNMENTS

12	19	54.3	6	13	AAR24624	Immunomodulatory p
13	19	54.3	6	13	AAR24690	Immunomodulatory p
14	19	54.3	6	19	AAW87231	Peptide determined
15	19	54.3	6	19	AAW87232	Peptide determined
16	19	54.3	6	19	AAW75344	Hexapeptide #9 bin
17	19	54.3	6	19	AAW75276	Hexapeptide #9 bin
18	19	54.3	6	22	AAG63041	Amino acid sequenc
19	19	54.3	7	13	AAR24651	Immunomodulatory p
20	19	54.3	7	13	AAR24652	Immunomodulatory p
21	19	54.3	7	13	AAR24664	Immunomodulatory p
22	19	54.3	7	17	AAR90639	S. solfatarius tr
23	18	51.4	5	17	AAW09684	Cyclic pentapeptid
24	18	51.4	5	17	AAB36778	HRG-beta1 variant
25	18	51.4	5	17	AAB36779	HRG-beta1 variant
26	18	51.4	7	20	AAY43592	GL04 synthetic pep
27	18	51.4	7	20	AAY47785	Amino acids 498-50
28	18	51.4	7	20	AAY49795	Elastolytic peptid
29	17	48.6	4	19	AAW47950	Elastolytic peptid
30	17	48.6	5	17	AAW05258	AE101 analogue eff
31	17	48.6	5	17	AAW02183	AE101 analogue eff
32	17	48.6	5	18	AAW12518	Cyclic pentapeptid
33	17	48.6	5	19	AAW47949	Residues 359-363 o
34	17	48.6	5	20	AAY39643	Interleukin-6 anta
35	17	48.6	5	21	AAB36777	Zea mays pathogen
36	17	48.6	5	21	AAB36775	HRG-beta1 variant
37	17	48.6	5	21	AAY93426	HRG-beta1 variant
38	17	48.6	5	21	AAY83328	Peptide motif from
39	17	48.6	5	11	AAR04740	Peptide motif of S
40	17	48.6	6	15	AAR56261	Deduced sequence a
41	17	48.6	6	19	AAW87227	Peptide correspond
42	17	48.6	6	19	AAW87228	Peptide determined
43	17	48.6	6	19	AAW87229	Peptide determined
44	17	48.6	6	19	AAW87230	Peptide determined
45	17	48.6	6	19	AAW47948	AE101 analogue eff
46	17	48.6	6	20	AAY41948	Rheumatoid arthrit
47	17	48.6	6	17	AAR85590	Human glycoprotein
48	17	48.6	6	19	AAW47946	AE101 analogue eff
49	17	48.6	7	20	AAY48887	Membrane dipeptida
50	17	48.6	7	20	AAW95149	Peptide ligand bin

```
XX
PT   New peptides, useful for generating canine autoantibodies against
PT   immunoglobulin E (IgE) to treat allergic symptoms -
XX
PS   Claim 5; Page 5; 38pp; English.
XX
CC   Sequences AAY69576-Y69586 and AAY69598-Y69599 represent peptides that may
CC   be used to induce anti-IgE (immunoglobulin E) autoantibodies in a dog.
CC   IgE is an important mediator of canine allergic responses, including Type
CC   I immediate hypersensitivity. It is estimated that up to 30% of all dogs
CC   suffer from allergies or allergy-related skin disorders, and allergic
CC   dermatitis is thought to affect between 3 and 15% of dogs. Common canine
CC   allergens include fleas, pollens, moulds and dust, with flea allergy
CC   being the most frequent. Once IgE has been released in response to
CC   allergen, the IgE binds to receptors on mast cells. The IgE molecules
CC   can then become cross-linked, which induces the mast cells to release
CC   histamine, which leads to the manifestation of allergic symptoms. The
CC   peptides of the invention correspond to epitopes of canine IgE, or
CC   conservative variations thereof, that are accessible in circulating IgE
CC   but not in IgE bound to mast cells. The autoantibodies produced on
CC   immunisation with such peptides therefore target only circulating IgE,
CC   avoiding the possibility that they may cross-link bound IgE molecules,
CC   which would exacerbate an ongoing allergic response even in the absence
CC   of allergen. Anti-IgE autoantibody induction results in an immune
CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targeting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog.
XX
SQ   Sequence   7 AA;

Query Match             100.0%;  Score 35;  DB 21;  Length 7;
Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
  Matches   7;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TLLEYRM 7
              |||||||
Db          1 tlleyrm 7

RESULT   2
AAY50878
ID   AAY50878 standard; peptide; 7 AA.
XX
AC   AAY50878;
XX
DT   24-FEB-2000  (first entry)
XX
DE   Antibody 8H.8 binding mimotope 1.
KW   Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW   epitope; prophylaxis; treatment; mimotope.
OS   Synthetic.
XX
XX
PN   EP957111-A2.
XX
PD   17-NOV-1999.
XX
PF   09-APR-1999;
XX
PR   09-APR-1998;   98US-0058331.
PR   30-MAR-1999;   99US-0281760.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-040833/04.
XX
PT   Binding proteins used for treatment or prophylaxis of canine allergy -
```

```
XX
PS   Claim 9; Page 12; 30pp; English.
XX
CC   This invention describes a novel binding protein which specifically
CC   binds to native canine free or B-cell bound IgE, and which doesn't bind
CC   to IgE when the IgE is bound to mast cells. THe peptide products of the
CC   invention have anti-allergic activity. The antibodies bind to defined
CC   epitopes on free or B-cell bound IgE molecules which have an important
CC   role in allergic reaction. The specific binding proteins are used to
CC   produce a pharmaceutical composition, preferably with a diluent, which
CC   can be used for prophylaxis or treatment of canine allergy. The
CC   AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC   invention.
XX
SQ   Sequence   7 AA;

Query Match             100.0%;  Score 35;  DB 21;  Length 7;
Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
Matches   7;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 TLLEYRM 7
              |||||||
Db          1 tlleyrm 7

RESULT   3
AAU03292
ID   AAU03292 standard; peptide; 7 AA.
XX
AC   AAU03292;
XX
DT   12-SEP-2001  (first entry)
XX
DE   Fruit fly G protein coupled receptors, DmGPCR6aL/bL ligand #41.
KW   Fruit fly; G protein coupled receptor; DmGPCR6aL/bL;
KW   human immunodeficiency virus; HIV; cancer; Parkinson's disease;
KW   diabetes; obesity; atherosclerosis; thrombosis; stroke; renal failure;
KW   inflammation; rheumatoid arthritis; autoimmune disorder;
KW   neurological disorder; schizophrenia; manic depression; dementia;
KW   severe mental retardation; dyskinesia; Huntington's disease;
KW   Tourette's syndrome; ligand.
OS   Drosophila melanogaster.
XX
XX
FH   Key             Location/Qualifiers
FT   Modified-site   7
FT                   /note= "C-terminus is amidated"
XX
PN   WO200131005-A2.
XX
PD   03-MAY-2001.
XX
PF   20-OCT-2000;   2000WO-US29002.
XX
PR   22-OCT-1999;   99US-0425676.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Lowery DE, Smith VG, Kubiak TA, Larsen MJ;
XX
DR   WPI; 2001-316333/33.
XX
PT   New Drosophila melanogaster GPCR nucleic acids and polypeptide useful
PT   for inducing an immune response, for identifying homologs and for
PT   treating e.g. diabetes, obesity and manic depression -
XX
PS   Example 9; Page 100; 110pp; English.
XX
CC   The sequence is a fruit fly G protein coupled receptors, DmGPCR6aL/bL,
CC   peptide ligand. The proteins are useful for inducing an immune response
CC   against itself in a mammal. The nucleic acids are useful for identifying
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:53:50 ; Search time 26.68 Seconds
(without alignments)
30.540 Million cell updates/sec

Title: us-09-281-760c-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 132282
Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database : A_Geneseq_1101:*
  1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
  2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
  3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
  4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
  5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
  6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
  7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
  8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
  9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
 10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
 11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
 12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
 13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
 14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
 15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
 16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
 17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
 18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
 19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
 20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
 21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
 22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	63	100.0	11	21	AAY69585	Canine IgE-derived
2	63	100.0	11	21	AAY50879	8H.8 binding canin
3	54	85.7	11	21	AAY50882	Antibody 8H.8 bind
4	40	63.5	11	21	AAY69591	Feline IgE-derived
5	40	63.5	11	21	AAY50881	Antibody 8H.8 bind
6	33	52.4	10	18	AAW22686	BSW17 peptide mime
7	30	47.6	7	21	AAY69590	Human IgE-derived
8	30	47.6	7	21	AAY50880	Antibody 8H.8 bind
9	30	47.6	9	17	AAW49652	Human leucocyte an
10	30	47.6	10	20	AAY40646	S2 derivative #8,
11	30	47.6	10	21	AAB29985	Scaffold protein S
12	29	46.0	7	19	AAW53900	Interleukin-1 rece
13	28	44.4	9	17	AAW49505	Human leucocyte an
14	27	42.9	6	18	AAW22699	BSW17 peptide mime
15	27	42.9	9	22	AAU15304	Schizophrenia-asso
16	27	42.9	10	18	AAW22688	BSW17 peptide mime
17	27	42.9	10	18	AAW22687	BSW17 peptide mime
18	26	41.3	8	15	AAR60424	Antiproliferative
19	26	41.3	8	15	AAR60447	Antiproliferative
20	26	41.3	10	22	AAB72239	Breast-cancer asso
21	26	41.3	11	17	AAW05276	PERB11-5D peptide.
22	26	41.3	11	19	AAW54288	Cytohesin 1 PH dom
23	25	39.7	7	20	AAY04349	Dengue antigenic P
24	25	39.7	8	15	AAR60425	Antiproliferative
25	25	39.7	8	15	AAR60446	Antiproliferative
26	25	39.7	8	20	AAR60431	Antiproliferative
27	25	39.7	9	20	AAY55436	HLA binding plu-1
28	25	39.7	9	20	AAY46859	Immunogenic peptid
29	25	39.7	10	17	AAW12590	Immunogenic peptid
30	25	39.7	10	20	AAY40648	S2 derivative #10,
31	25	39.7	10	20	AAY40649	S2 binding peptid
32	25	39.7	10	20	AAY40650	S2 derivative #11,
33	25	39.7	10	21	AAB29987	Scaffold protein S
34	25	39.7	10	21	AAB29988	Scaffold protein S
35	25	39.7	10	21	AAB29989	Scaffold protein S
36	24	38.1	6	17	AAR90506	Hybridoma ATCC HB-
37	24	38.1	6	20	AAY55233	ATCC HB 11884 mono
38	24	38.1	6	20	AAW95714	Peptide derived fr
39	24	38.1	6	21	AAY86979	Human haematopoiet
40	24	38.1	6	22	AAB69385	Grand fir monoterp
41	24	38.1	8	14	AAR36082	Hepatitis C virus
42	24	38.1	8	14	AAR36083	Hepatitis C virus
43	24	38.1	8	15	AAR60437	Antiproliferative
44	24	38.1	8	20	AAY03896	Predicted binding
45	24	38.1	8	22	AAJ00475	Hepatitis C virus
46	24	38.1	8	22	AAJ00972	Hepatitis C virus
47	24	38.1	8	22	AAJ01677	Hepatitis C virus
48	24	38.1	8	22	AAB78520	HIV-1 gp 41 enhanc
49	24	38.1	9	18	AAW45164	Human cartilage gl
50	24	38.1	9	18	AAW37010	TRP-2 derived pote

RESULT 1
AAY69585

ID   AAY69585 standard; peptide; 11 AA.
XX
AC   AAY69585;
XX
DT   19-APR-2000 (first entry)
XX
DE   Canine IgE-derived peptide E3a.5, SEQ ID NO:10.
XX
KW   Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW   allergy; canine; prophylaxis; treatment.
XX
OS   Canis familiaris.
XX
PN   EP955311-A2.
XX
PD   10-NOV-1999.
XX
PF   09-APR-1999; 99EP-0107020.
XX
PR   09-APR-1998; 98US-0058332.
PR   30-MAR-1999; 99US-0281761.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-118182/11.

ALIGNMENTS

XX
PT New peptides, useful for generating canine autoantibodies against
    immunoglobulin E (IgE) to treat allergic symptoms -
XX
PS Claim 5; Page 5; 38pp; English.
XX
CC Sequences AAY69576-Y69586 and AAY69598-Y69599 represent peptides that may
CC be used to induce anti-IgE (immunoglobulin E) autoantibodies in a dog.
CC IgE is an important mediator of canine allergic responses, including Type
CC I immediate hypersensitivity. It is estimated that up to 30% of all dogs
CC suffer from allergies or allergy-related skin disorders, and allergic
CC dermatitis is thought to affect between 3 and 15% of dogs. Common canine
CC allergens include fleas, pollens, moulds and dust, with flea allergy
CC being the most frequent. Once IgE has been released in response to
CC allergen, the IgE binds to receptors on mast cells. The IgE molecules
CC can then become cross-linked, which induces the mast cells to release
CC histamine, which leads to the manifestation of allergic symptoms. The
CC peptides of the invention correspond to epitopes of canine IgE, or
CC conservative variations thereof, that are accessible in circulating IgE
CC but not in IgE bound to mast cells. The autoantibodies produced on
CC immunisation with such peptides therefore target only circulating IgE,
CC avoiding the possibility that they may cross-link bound IgE molecules,
CC which would exacerbate an ongoing allergic response even in the absence
CC of allergen. Anti-IgE autoantibody induction results in an immune
CC response that targets the IgE-expressing B cell, which affects IgE
CC synthesis by downregulating IgE production by B cells and/or targeting
CC the IgE-producing B cell for destruction. The peptides of the invention
CC are therefore useful for the treatment or prophylaxis of allergic
CC symptoms in a dog.
XX
SQ Sequence 11 AA;
XX
.
Query Match   100.0%; Score 63; DB 21; Length 11;
Best Local Similarity   100.0%; Pred. No. 7.9e-05;
Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
.
Qy   1 GMNLTWYRESK 11
              |||||||||||
Db   1 gmnltwyresk 11

RESULT 2
ID AAY50879
AC AAY50879;
XX
ID AAY50879 standard; peptide; 11 AA.
XX
AC AAY50879;
XX
DT 24-FEB-2000 (first entry)
XX
DE 8H.8 binding canine mimotope 1.
XX
KW Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -

XX
PS Claim 10; Page 13; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 11 AA;
XX
Query Match   100.0%; Score 63; DB 21; Length 11;
Best Local Similarity   100.0%; Pred. No. 7.9e-05;
Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;
.
Qy   1 GMNLTWYRESK 11
              |||||||||||
Db   1 gmnltwyresk 11

RESULT 3
ID AAY50882
AC AAY50882;
XX
ID AAY50882 standard; peptide; 11 AA.
XX
AC AAY50882;
XX
DT 24-FEB-2000 (first entry)
XX
DE Antibody 8H.8 binding substituted canine mimotope 1.
XX
KW Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW epitope; prophylaxis; treatment; mimotope.
XX
OS Synthetic.
XX
PN EP957111-A2.
XX
PD 17-NOV-1999.
XX
PF 09-APR-1999; 99EP-0107035.
XX
PR 09-APR-1998; 98US-0058331.
PR 30-MAR-1999; 99US-0281760.
XX
PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-040833/04.
XX
PT Binding proteins used for treatment or prophylaxis of canine allergy -
XX
PS Example 2; Page 13; 30pp; English.
XX
CC This invention describes a novel binding protein which specifically
CC binds to native canine free or B-cell bound IgE, and which doesn't bind
CC to IgE when the IgE is bound to mast cells. The peptide products of the
CC invention have anti-allergic activity. The antibodies bind to defined
CC epitopes on free or B-cell bound IgE molecules which have an important
CC role in allergic reaction. The specific binding proteins are used to
CC produce a pharmaceutical composition, preferably with a diluent, which
CC can be used for prophylaxis or treatment of canine allergy.
CC AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC invention.
XX
SQ Sequence 11 AA;

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:16:05 ; Search time 85.5 Seconds
                (without alignments)
                2.144 Million cell updates/sec

Title: US-09-281-760C-1
Perfect score: 18
Sequence: 1 LXXYR 5

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Database :   SwissProt_39:*

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

1mun DB seq length: 0
1mun DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	AC	Description
1	16	88.9	38	1	CPRP_CANPG	P81033	cancer pagu
2	16	88.9	69	1	YOG6_CAEEL	P34615	caenorhabdi
3	16	88.9	93	1	Y292_METJA	Q57740	methanococc
4	16	88.9	99	1	SYO2_HUMAN	P13500	homo sapien
5	16	88.9	99	1	SYO2_MACFA	Q9myn4	macaca fasc
6	16	88.9	106	1	DCMA_METTE	Q08368	methanosarc
7	16	88.9	109	1	YIR1_YEAST	P40440	saccharomyc
8	16	88.9	113	1	KV2B_HUMAN	P01615	homo sapien
9	16	88.9	191	1	YGL2_STRCO	P40182	streptomyce
10	16	88.9	195	1	CNPF_CHICK	O02011	gallus gall
11	16	88.9	199	1	TDX2_THEAC	Q9hj13	thermoplasm
12	16	88.9	199	1	TDX_TRYBR	Q26695	trypanosoma
13	16	88.9	200	1	Y755_SYNY3	O55624	synechocyst
14	16	88.9	202	1	KAPS_YEAST	O00196	saccharomyc
15	16	88.9	207	1	CLI3_HUMAN	O95833	homo sapien
16	16	88.9	211	1	KAD_TREPA	O83604	treponema p
17	16	88.9	215	1	YF52_HAEIN	P44251	haemophilus
18	16	88.9	217	1	UNG_STRAG	Q9xd58	streptococc
19	16	88.9	220	1	UNG_BUCAI	P57280	buchnera ap
20	16	88.9	222	1	RS3A_METJA	P54059	methanococc
21	16	88.9	228	1	VIF_OMVVS	P16902	ovine lenti
22	16	88.9	230	1	VIF_VILV	P03403	visna lenti
23	16	88.9	230	1	VIF_VILV1	P23430	visna lenti
24	16	88.9	230	1	YDB3_YEAST	Q04004	saccharomyc
25	16	88.9	231	1	VIC1_AGRRH	P07165	agrobacteri
26	16	88.9	231	1	VIC1_AGRT5	P07165	agrobacteri
27	16	88.9	231	1	VIC1_AGRT6	P07165	agrobacteri
28	16	88.9	232	1	YFP_ECOLI	P06665	escherichia
29	16	88.9	237	1	GSP2_PSEAE	Q47319	pseudomonas
30	16	88.9	238	1	DHSB_ECOLI	O00517	escherichia
31	16	88.9	247	1	CORK_PSEDE	P07014	pseudomonas
32	16	88.9	254	1	PMM_YEAST	P21920	saccharomyc
33	16	88.9	256	1	TPS1_BRARE	Q9pte6	brachydanio
34	16	88.9	257	1	PMM_SCHPO	Q9tj2	schizosacch
35	16	88.9	260	1	CY12_SOLTU	P29610	solanum tub
36	16	88.9	260	1	Y41O_RHISN	P55555	rhizobium s
37	16	88.9	261	1	TFPP_MORBO	P20666	moraxella b
38	16	88.9	262	1	GP6R_CHLMU	Q9pif7	chlamydia m
39	16	88.9	271	1	CARL_MOUSE	Q08331	mus musculu
40	16	88.9	271	1	CARI_RAT	P47728	rattus norv
41	16	88.9	276	1	PYG1_SYNEL	P50039	synechococc
42	16	88.9	277	1	MSMG_STRMU	Q00751	streptococc
43	16	88.9	280	1	SPEE_PYRHO	O57950	pyrococcus
44	16	88.9	284	1	HUPQ_AZOCH	Q43957	azotobacter
45	16	88.9	289	1	Y097_MYCTU	O10893	mycobacteri
46	16	88.9	297	1	SMF_BACSU	P39813	bacillus su
47	16	88.9	297	1	YN24_CAEEL	P45964	caenorhabdi
48	16	88.9	298	1	XERD_SALTY	P55889	salmonella
49	16	88.9	308	1	Y705_DEIRA	Q9rvg3	deinococcus
50	16	88.9	309	1	URIC_ARATH	O04420	arabidopsis

ALIGNMENTS

RESULT 1
CPRP_CANPG

```
ID   CPRP_CANPG     STANDARD;      PRT;    38 AA.
AC   P81033;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CHH PRECURSOR RELATED PEPTIDE (CPRP).
OS   Cancer pagurus (Rock crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Cancroidea; Cancridae; Cancer.
OX   NCBI_TaxID=6755;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Sinus gland;
RX   MEDLINE=99025664; PubMed=9809792;
RA   Chung J.S., Wilkinson M.C., Webster S.G.;
RT   "Amino acid sequences of both isoforms of crustacean hyperglycemic
RT   hormone (CHH) and corresponding precursor-related peptide in Cancer
RT   pagurus.";
RL   Regul. Pept. 77:17-24(1998).
CC   -!- TISSUE SPECIFICITY: PRODUCED BY THE MEDULLA TERMINALIS X-ORGAN IN
CC       THE EYESTALKS AND TRANSPORTED TO THE SINUS GLAND WHERE IT IS
CC       STORED AND RELEASED.
KW   Neuropeptide; Hormone.
SQ   SEQUENCE   38 AA;  3969 MW;  C979C87EE31ABB90 CRC64;
```

```
Query Match             88.9%;  Score 16;  DB 1;  Length 38;
Best Local Similarity   60.0%;  Pred. No. 1.2e+02;
Matches    3;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LXXYR 5
              | ||
Db         13 LASYR 17
```

RESULT 2
YOG6_CAEEL

```
ID   YOG6_CAEEL     STANDARD;      PRT;    69 AA.
AC   P34615;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-FEB-1994 (Rel. 28, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   HYPOTHETICAL 7.9 KDA PROTEIN ZK112.6 IN CHROMOSOME III.
GN   ZK112.6.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
```

RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA   Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA   Johnston L., Jones M., Kershaw J., Kirsten J., Laisster N.,
RA   Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showkeen R.,
RA   Sims M., Smaldon N., Smith A., Smith M., Sonnhammer E., Staden R.,
RA   Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA   Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA   Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).

DR   EMBL; L14324; AAA28183.1; -.
DR   PIR; S44888; S44888.
DR   WormPep; ZK112.6; CE00377.
KW   Hypothetical protein.
SQ   SEQUENCE 69 AA; 7886 MW; 348D05DF5EEBA54A CRC64;

Qy   Query Match   88.9%; Score 16; DB 1; Length 69;
.   Best Local Similarity 60.0%; Pred. No. 2.1e+02;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXYR 5
           | ||
Db   8 LAAYR 12

RESULT 3
Y292_METJA
ID   Y292_METJA   STANDARD;   PRT;   93 AA.
AC   Q57740;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL PROTEIN MJ0292.
GN   MJ0292.
OS   Methanococcus jannaschii.
OC   Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC   Methanococcus.
OX   NCBI_TaxID=2190;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DAL-1 / DSM 2661 / ATCC 43067;
RX   MEDLINE=96637999; PubMed=8688087;
RA   Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA   Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA   Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
RA   Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA   Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA   Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA   Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA   Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT   "Complete genome sequence of the methanogenic archaeon, Methanococcus
RT   jannaschii.";
RL   Science 273:1058-1073(1996).
CC   -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).
CC   ---

DR   EMBL; U67484; AAB98277.1; -.
DR   TIGR; MJ0292; -.
KW   Hypothetical protein; Transmembrane; Complete proteome.
FT   TRANSMEM   8   28   POTENTIAL.
FT   TRANSMEM   54   74   POTENTIAL.
FT   TRANSMEM   75   86   ALA-RICH.
FT   DOMAIN   86   93   ALA-RICH.
SQ   SEQUENCE 93 AA; 10482 MW; B2A8C8I15AC160C CRC64;

Qy   Query Match   88.9%; Score 16; DB 1; Length 93;
.   Best Local Similarity 60.0%; Pred. No. 2.9e+02;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXYR 5
           | ||
Db   26 LASYR 30

RESULT 4
SY02_HUMAN
ID   SY02_HUMAN   STANDARD;   PRT;   99 AA.
AC   P13500; O9UDF3;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-JAN-1990 (Rel. 13, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   SMALL INDUCIBLE CYTOKINE A2 PRECURSOR (MONOCYTE CHEMOTACTIC PROTEIN 1)
DE   (MCP-1) (MONOCYTE CHEMOATTRACTANT PROTEIN-1) (MONOCYTE CHEMOTACTIC
DE   AND ACTIVATING FACTOR) (MCAF) (MONOCYTE SECRETORY PROTEIN JE) (HC11).
GN   SCYA2 OR MCP1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89165862; PubMed=2923622;
RA   Furutani Y., Nomura H., Notake M., Oyamada Y., Fukui T., Yamada M.,
RA   Larsen C.G., Oppenheim J.J., Matsushima K.;
RT   "Cloning and sequencing of the cDNA for human monocyte chemotactic
RT   and activating factor (MCAF).";
RL   Biochem. Biophys. Res. Commun. 159:249-255(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90097880; PubMed=2513477;
RA   Rollins B.J., Stier P., Ernst T., Wong G.G.;
RT   "The human homolog of the JE gene encodes a monocyte secretory
RT   protein.";
RL   Mol. Cell. Biol. 9:4687-4695(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89153605; PubMed=2465924;
RA   Yoshimura T., Yuhki N., Moore S.K., Appella E., Lerman M.I.,
RA   Leonard E.J.;
RT   "Human monocyte chemoattractant protein-1 (MCP-1). Full-length cDNA
RT   cloning, expression in mitogen-stimulated blood mononuclear
RT   leukocytes, and sequence similarity to mouse competence gene JE.";
RL   FEBS Lett. 244:487-493(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90290466; PubMed=2357211;
RA   Shyy Y.J., Li Y.S., Kolattukudy P.E.;
RT   "Structure of human monocyte chemotactic protein gene and its
RT   regulation by TPA.";
RL   Biochem. Biophys. Res. Commun. 169:346-351(1990).
RN   [5]

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:11:15 ; Search time 102.84 Seconds
                 (without alignments)
                 3.704 Million cell updates/sec

Title: US-09-281-760C-1
Perfect score: 18
Sequence: 1 LXXYR 5

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Database :   PIR 68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Total number of hits satisfying chosen parameters: 219241

imum DB seq length: 0
ximum DB seq length: 2000000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	88.9	69	2	S44888	ZK112.6 protein -
2	16	88.9	74	2	T13344	hypothetical prote
3	16	88.9	92	2	T09146	late-embryogenesis
4	16	88.9	93	2	I47024	Ig heavy chain v-v
5	16	88.9	93	2	E64336	conserved hypothet
6	16	88.9	99	2	A60299	monocyte chemoattr
7	16	88.9	101	2	T37064	hypothetical prote
8	16	88.9	101	2	D82808	hypothetical prote
9	16	88.9	105	2	E69058	hypothetical prote
10	16	88.9	105	2	H72533	hypothetical prote
11	16	88.9	105	2	A48868	carbon-monoxide de
12	16	88.9	106	2	S50356	sugar transport pr
13	16	88.9	109	2	B72615	hypothetical prote
14	16	88.9	109	2	K2HUFR	Ig kappa chain v-I
15	16	88.9	113	2	S22559	Ig heavy chain V r
16	16	88.9	114	2	T41216	very hypothetical
17	16	88.9	117	2	A60894	hypothetical prote
18	16	88.9	130	2	A35946	glial fibrillary a
19	16	88.9	130	2	T44809	hypothetical prote
20	16	88.9	132	2	F70650	hypothetical prote
21	16	88.9	140	2	H75426	conserved hypothet
22	16	88.9	146	2	E75529	hypothetical prote
23	16	88.9	148	2	C71263	probable phenylace
24	16	88.9	149	2	H70638	probable flagellar
25	16	88.9	151	2	H75298	probable thiol-spe
26	16	88.9	152	2	B72690	hypothetical prote
27	16	88.9	156	2	D84499	hypothetical prote
28	16	88.9	161	2	T35369	hypothetical prote
29	16	88.9	161	2	G83914	hypothetical prote
30	16	88.9	165	2	T01421	pistil extensin-li
31	16	88.9	166	2	B83700	hypothetical prote
32	16	88.9	168	2	E83437	hypothetical prote
33	16	88.9	170	2	S76417	hypothetical prote
34	16	88.9	175	2	T15067	hypothetical prote
35	16	88.9	176	2	G84214	hypothetical prote
36	16	88.9	181	2	C84233	hypothetical prote
37	16	88.9	190	2	S41476	calretinin - mouse
38	16	88.9	191	2	T35508	hypothetical prote
39	16	88.9	195	2	JH0680	ciliary neurotroph
40	16	88.9	200	2	S76284	hypothetical prote
41	16	88.9	202	2	S17244	adenylylsulfate ki
42	16	88.9	206	2	T22345	hypothetical prote
43	16	88.9	207	2	B27626	probable lipoprote
44	16	88.9	210	2	S27489	hypothetical prote
45	16	88.9	210	2	T22344	probable adenylate
46	16	88.9	211	2	D71306	probable acyl carr
47	16	88.9	212	2	H83547	probable adenylate
48	16	88.9	215	2	I64035	hypothetical prote
49	16	88.9	219	2	B81908	probable uracil-DN
50	16	88.9	219	2	H81107	uracil-DNA glycosy

ALIGNMENTS

RESULT 1
S44888
ZK112.6 protein - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 14-Sep-1994 #sequence_revision 12-May-1995 #text_change 23-Mar-2001
C;Accession: S44888
R;Du, Z.
submitted to the EMBL Data Library, May 1993
A;Description: Sequence of the C. elegans cosmid ZK112.
A;Reference number: S44616
A;Accession: S44888
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-69 <DUZ>
A;Cross-references: EMBL:L14324; NID:g289740; PID:g289743
C;Genetics:
A;Introns: 20/3

Query Match   88.9%;   Score 16;   DB 2;   Length 69;
Best Local Similarity   60.0%;
Matches   3;   Conservative   0;   Pred. No. 5.2e+02;
Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 LXXYR 5
        | ||
Db   8 LAAYR 12

RESULT 2
T13344
hypothetical protein 55 - Streptococcus phage phi-O1205
C;Species: Streptococcus phage phi-O1205
C;Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 24-Sep-1999
C;Accession: T13344
R;Stanley, E.; Fitzgerald, G.F.; Le Marrec, C.; Fayard, B.; van Sinderen, D.
Microbiology 143, 3417-3429, 1997
A;Title: Sequence analysis and characterization of phi O1205, a temperate bacteriopha
A;Reference number: Z17654; MUID:98048466
A;Accession: T13344
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-74 <STA>
A;Cross-references: EMBL:U88974; NID:g2444080; PID:g2444135; PIDN:AAC79571.1
A;Experimental source: host Streptococcus thermophilus strain CNRZ1205

Query Match   88.9%;   Score 16;   DB 2;   Length 74;

Best Local Similarity   60.0%;   Pred. No. 5.5e+02;   Mismatches 2;   Indels 0;   Gaps 0;
Matches   3;   Conservative   0;

Qy   1 LXXYR   5
           | ||
Db   59 LSAYR 63

RESULT 3
T09146
late-embryogenesis protein EMB11 - white spruce
C;Species: Picea glauca (white spruce)
C;Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C;Accession: T09146
R;Dong, J.Z.; Dunstan, D.I.
A;Title: Gene expression during somatic embryogenesis.
A;Description: Gene expression during somatic embryogenesis.
submitted to the EMBL Data Library, June 1996
R;Reference number: Z16588
A;Accession: T09146
C;Genetics:
C;Cross-references: EMBL:L47628; NID:g1350491; PID:g1350492
A;Molecule type: mRNA
A;Residues: 1-92 <DON>
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Gene: EMB11

Query Match   88.9%;   Score 16;   DB 2;   Length 92;
Best Local Similarity   60.0%;   Pred. No. 6.8e+02;
Matches   3;   Conservative   0;   Mismatches 2;   Indels 0;   Gaps 0;

Qy   1 LXXYR   5
           | ||
Db   6 LSAYR 10

RESULT 4
I47624
Ig heavy chain V-VIII region - African clawed frog (fragment)
C;Species: Xenopus laevis (African clawed frog)
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 11-Jan-2000
C;Accession: I47624
R;Haire, R.N.; Amemiya, C.T.; Suzuki, D.; Litman, G.W.
J. Exp. Med. 171, 1721-1737, 1990
A;Title: Eleven distinct V-H gene families and additional patterns of sequence variation
A;Reference number: A47624; MUID:90237760
A;Accession: I47624
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-93 <HAI>
A;Superfamily: immunoglobulin V region; immunoglobulin homology
C;Keywords: heterotetramer; immunoglobulin

Query Match   88.9%;   Score 16;   DB 2;   Length 93;
Best Local Similarity   60.0%;   Pred. No. 6.9e+02;
Matches   3;   Conservative   0;   Mismatches 2;   Indels 0;   Gaps 0;

Qy   1 LXXYR   5
           | ||
Db   47 LAAYR 51

RESULT 5
E64336
conserved hypothetical protein MJ0292 - Methanococcus jannaschii
C;Species: Methanococcus jannaschii
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
C;Accession: E64336
R;Bult, C.J.; White, O.; Olson, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, A.;
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;
; ison, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
Science 273, 1058-1073, 1996

A;Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
A;Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc
A;Reference number: A64300; MUID:96337999
A;Accession: E64336
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-93 <BUL>
A;Cross-references: GB:U67484; GB:L77117; NID:g1591009; PIDN:AAB98277.1; PID:g1591014
C;Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0292

Query Match   88.9%;   Score 16;   DB 2;   Length 93;
Best Local Similarity   60.0%;   Pred. No. 6.9e+02;
Matches   3;   Conservative   0;   Mismatches 2;   Indels 0;   Gaps 0;

Qy   1 LXXYR   5
           | ||
Db   26 LASYR 30

RESULT 6
A60299
monocyte chemoattractant protein 1 precursor - human
A;Alternate names: GDCF-1; glioma-derived monocyte chemotactic factor 1; MCAF; MCP-1;
N;Contains: glioma-derived chemotactic factor 2 (GDCF-2)
R;Rollins, B.J.; Stier, P.; Ernst, T.; Wong, G.G.
C;Species: Homo sapiens (man)
C;Date: 20-Feb-1993 #sequence_revision 20-Feb-1993 #text_change 16-Jul-1999
C;Accession: A35474; A33476; S03339; I51841; A60299; A32300; A33396; A34561; I57488;
R;Shyy, Y.J.; Li, Y.S.; Kolattukudy, P.E.
Biochem. Biophys. Res. Commun. 169, 346-351, 1990
A;Title: Structure of human monocyte chemotactic protein gene and its regulation by T
A;Reference number: A35474; MUID:90204466
A;Accession: A35474
A;Molecule type: DNA
A;Residues: 1-99 <SHY>
A;Cross-references: GB:M37719; NID:g187447; PIDN:AAA18102.1; PID:g487124
Mol. Cell. Biol. 9, 4687-4695, 1989
A;Title: The human homolog of the JE gene encodes a monocyte secretory protein.
A;Reference number: A33476; MUID:90097880
A;Accession: A33476
A;Molecule type: mRNA
A;Residues: 1-99 <ROL>
A;Cross-references: GB:M30816; GB:M31625; GB:M31626; NID:g188701; PIDN:AAA36330.1; PI
R;Yoshimura, T.; Yuhki, N.; Moore, S.K.; Appella, E.; Lerman, M.I.; Leonard, E.J.
FEBS Lett. 244, 487-493, 1989
A;Title: Human monocyte chemoattractant protein-1 (MCP-1). Full-length cDNA cloning,
;Reference number: S03339; MUID:89153605
A;Accession: S03339
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-99 <YOS>
A;Cross-references: GB:X14768; NID:g34513; PIDN:CAA32876.1; PID:g34514
A;Experimental source: glioma cell line U-105MG
R;Yoshimura, T.; Leonard, E.J.
Adv. Exp. Med. Biol. 305, 47-56, 1991
A;Title: Human monocyte chemoattractant protein-1 (MCP-1).
A;Reference number: I51841; MUID:92095166
A;Accession: I51841
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-99 <YO2>
A;Cross-references: GB:S71513; NID:g240867; PIDN:AAB20651.1; PID:g240868
R;Bottazzi, B.; Colotta, F.; Sica, A.; Nobili, N.; Mantovani, A.
Int. J. Cancer 45, 795-797, 1990
A;Title: A chemoattractant expressed in human sarcoma cells (tumor-derived chemotacti
-1/MCAF).
A;Reference number: A60299; MUID:90216082
A;Accession: A60299

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:14:20 ; Search time 191.29 Seconds
                                              (without alignments)
                                              3.823 Million cell updates/sec

Title: US-09-281-760C-1
Perfect score: 18
Sequence: 1 LXXYR 5
Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 473505
Listing first 50 summaries

Searched:

Post-processing: Minimum Match 0%
                 Maximum Match 100%

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   SPTREMBL_17:*
                 1: sp_archea:*
                 2: sp_bacteria:*
                 3: sp_fungi:*
                 4: sp_human:*
                 5: sp_invertebrate:*
                 6: sp_mammal:*
                 7: sp_mhc:*
                 8: sp_organelle:*
                 9: sp_phage:*
                 10: sp_plant:*
                 11: sp_rodent:*
                 12: sp_virus:*
                 13: sp_vertebrate:*
                 14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	DB ID	Description
1	16	88.9	44	6	Q9BG83	pongo pygma
2	16	88.9	59	6	O97749	bos taurus
3	16	88.9	74	2	O34086	streptococc
4	16	88.9	74	6	O21986	streptococc
5	16	88.9	78	1	Q9V2Y7	methanobact
6	16	88.9	84	5	P91813	toxocara ca
7	16	88.9	90	2	O50475	proteus vul
8	16	88.9	91	12	Q72831	human immun
9	16	88.9	92	10	O40840	picea glauc
10	16	88.9	94	2	Q9MC42	bacteriopha
11	16	88.9	97	6	O77661	papio anubi
12	16	88.9	101	2	Q9S1Z2	streptomyce
13	16	88.9	101	2	Q9PG85	xylella fas
14	16	88.9	102	6	Q9AYW1	lactococcus
15	16	88.9	102	13	Q9DDQ7	decapterus
16	16	88.9	102	13	Q9DDQ5	decapterus
17	16	88.9	102	13	Q9DDQ4	decapterus
18	16	88.9	102	13	Q9DDQ3	decapterus
19	16	88.9	102	13	Q9DDQ1	decapterus
20	16	88.9	102	13	Q9DDQ0	decapterus
21	16	88.9	102	13	Q9DDP9	decapterus
22	16	88.9	102	13	Q9DD71	decapterus
23	16	88.9	103	2	Q9KW07	pseudomonas
24	16	88.9	105	1	O27488	methanobact
25	16	88.9	105	1	Q9Y9P4	aeropyrum p
26	16	88.9	109	1	Q9Y9P4	aeropyrum p
27	16	88.9	110	12	Q9YC72	heliocoverp
28	16	88.9	112	5	Q99GX6	leishmania
29	16	88.9	112	2	Q9AJZ8	streptomyce
30	16	88.9	117	3	Q9Y7P7	schizosacch
31	16	88.9	118	2	Q9L5M8	salmonella
32	16	88.9	121	8	Q9TAP8	heterophyes
33	16	88.9	126	4	O14956	homo sapien
34	16	88.9	128	1	Q9R2C0	rattus norv
35	16	88.9	130	2	Q9HHI0	halobacteri
36	16	88.9	130	2	O34506	bacillus su
37	16	88.9	130	2	Q9H918	bacillus su
38	16	88.9	130	2	Q9L7W8	bacillus su
39	16	88.9	130	10	Q9M4G0	arabidopsis
40	16	88.9	132	2	P95089	mycobacteri
41	16	88.9	135	12	O68190	hepatitis c
42	16	88.9	139	2	Q9RDQ0	streptomyce
43	16	88.9	140	2	Q9RV53	deinococcus
44	16	88.9	141	2	Q9FBM4	carboxydoth
45	16	88.9	143	10	Q9XFV8	chlamydomon
46	16	88.9	144	2	Q9C866	mycobacteri
47	16	88.9	144	2	Q99GG4	streptococc
48	16	88.9	146	2	Q9RS06	deinococcus
49	16	88.9	148	2	O83913	treponema p
50	16	88.9	149	2	P95258	mycobacteri

ALIGNMENTS

RESULT 1
ID   Q9BG83
AC   Q9BG83;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   MONOCYTE CHEMOTACTIC PROTEIN 1 (FRAGMENT).
GN   MCP1.
OS   Pongo pygmaeus (Orangutan).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pongo.
OX   NCBI_TaxID=9600;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Ayi T.C., Chew X.L.A., Yap P.H.E.;
RT   "Comparison between human and primate genomes.";
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF324469; AAK13442.1; -.
FT   NON_TER   1
FT   NON_TER   44
SQ   SEQUENCE   44 AA;   4830 MW;   1DE3C10EB83CD190 CRC64;

            Query Match   88.9%; Score 16; DB 6; Length 44;
            Best Local Similarity   60.0%;
            Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

            Qy   1 LXXYR 5
                          | ||
            Db   40 LASYR 44

RESULT 2
ID   O97749
AC   O97749;   PRELIMINARY;   PRT;   59 AA.

DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE   BOVINE PROTHROMBIN MRNA, PARTIAL CDS (FRAGMENT).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=8077773; PubMed=3000440;
RA   Irwin D.M., Ahern K.G., Pearson G.D.N., Macgillivray R.T.A.;
RT   "Characterization of the bovine prothrombin gene.";
RL   Biochemistry 24:6854-6861(1985).
DR   EMBL; M12049; AAA30780.1; -.
FT   NON_TER   1
FT   NON_TER   59
Q   SEQUENCE   59 AA; 6641 MW; 5570708FAE94A74A CRC64;

Qy   1 LXXYR 5
         | |
Db   17 LTSYR 21

Query Match   88.9%; Score 16; DB 6; Length 59;
Best Local Similarity   60.0%; Pred. No. 1.1e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

RESULT   3
ID   O34086   PRELIMINARY;   PRT;   74 AA.
AC   O34086;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   ORF55.
OS   Streptococcus thermophilus.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Streptococcus.
OX   NCBI_TaxID=1308;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=CNR21205;
RA   Stanley E., Fitzgerald G.F., Le Marrec C., Fayard B., van Sinderen D.;
RT   "Sequence analysis and characterization of phi O1205, a temperate
RT   bacteriophage infecting Streptococcus thermophilus CNRZ1205.";
RL   Microbiology 143:0-0(0).
DR   EMBL; U88974; AAC79571.1; -.
SQ   SEQUENCE   74 AA; 8609 MW; 141IFE219D50BED1 CRC64;

Qy   1 LXXYR 5
         | |
Db   59 LSAYR 63

Query Match   60.0%; Pred. No. 1.5e+03;
Best Local Similarity   60.0%; Score 16; DB 2; Length 74;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

RESULT   4
ID   O21986   PRELIMINARY;   PRT;   74 AA.
AC   O21986;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   ORF74A PROTEIN.
OS   Streptococcus thermophilus bacteriophage Sfi21.
OC   Viruses; dsDNA viruses, no RNA stage; Tailed phages; Siphoviridae.

OX   NCBI_TaxID=64186;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96204576; PubMed=8623559;
RA   Bruttin A., Brussow H.;
RT   "Site-specific spontaneous deletions in three genome regions of a
RT   temperate Streptococcus thermophilus phage.";
RL   Virology 219:96-104(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97348856; PubMed=9201223;
RA   Bruttin A., Desiere F., Lucchini S., Foley S., Brussow H.;
RT   "Characterization of the lysogeny DNA module from the temperate
RT   Streptococcus thermophilus bacteriophage phi Sfi21.";
RL   Virology 233:136-148(1997).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98160788; PubMed=9499809;
RA   Desiere F., Lucchini S., Brussow H.;
RT   "Evolution of Streptococcus thermophilus bacteriophage genomes by
RT   modular exchanges followed by point mutations and small deletions and
RT   insertions.";
RL   Virology 241:345-356(1998).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Desiere F., Lucchini S., Bruttin A., Brussow H.;
RT   "Streptococcus thermophilus bacteriophage Sfi21 complete genome.";
RL   Submitted (DEC-1998) to the EMBL/GenBank/DDBJ databases.
RL   EMBL; X95646; CAA64946.1; -.
DR   EMBL; AF115103; AAD44092.1; -.
SQ   SEQUENCE   74 AA; 8592 MW; 12D73E275B99E5E3F CRC64;

Qy   1 LXXYR 5
         | |
Db   59 LSAYR 63

Query Match   88.9%; Score 16; DB 9; Length 74;
Best Local Similarity   60.0%; Pred. No. 1.5e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

RESULT   5
ID   Q9V2Y7   PRELIMINARY;   PRT;   78 AA.
AC   Q9V2Y7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   TRANSMEMBRANE PROTEIN 8.3 KDA.
OS   Methanobacterium thermoautotrophicum (strain Marburg / DSM 2133).
OC   Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;
OC   Methanothermobacter.
OX   NCBI_TaxID=79929;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MARBURG;
RA   Tersteegen A., Hedderich R.;
RT   "Methanobacterium thermoautotrophicum encodes two multisubunit
RT   membrane-bound [NiFe] hydrogenases.Transcription of the operons and
RT   sequence analysis of the deduced proteins.";
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ243656; CAB52781.1; -.
KW   Transmembrane.
SQ   SEQUENCE   78 AA; 8285 MW; 1CEDDF8DC27F63D7 CRC64;

Qy   1 LXXYR 5
         | |

Query Match   88.9%; Score 16; DB 1; Length 78;
Best Local Similarity   60.0%; Pred. No. 1.5e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

```
OM protein - protein search, using sw model

              GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:           April 21, 2002, 10:11:10 ; Search time 158.02 Seconds
                                              (without alignments)
                                              2.344 Million cell updates/sec

Title:            US-09-281-760C-1
Perfect score:    18
Sequence:         1 LXXYR 5

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:       522463

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 50 summaries
Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :   A_Geneseq_1101:*
        1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
        2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
        3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
        4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
        5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
        6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
        7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
        8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
        9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
       10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
       11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
       12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
       13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
       14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
       15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
       16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
       17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
       18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
       19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
       20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
       21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
       22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

```
                   %       Query
Result           Match   Length  DB
 NO.   Score    Length    DB     ID         Description
-----  -----    ------   ------  --         -----------
  1     16      88.9       7     20  AAY34076  Human monocyte che
  2     16      88.9       7     20  AAY34080  Human monocyte che
  3     16      88.9       7     20  AAY34084  Human monocyte che
  4     16      88.9      10     22  AAB05375  Glial fibrillary a
  5     16      88.9      11     12  AAR13244  Nocifensin anti-in
  6     16      88.9      11     15  AAR48303  Antiinflammatory p
  7     16      88.9      11     15  AAR48304  Antiinflammatory p
  8     16      88.9      11     15  AAR48309  Antiinflammatory p
  9     16      88.9      23     21  AAB15851  Chemokine derived
 10     16      88.9      23     22  AAY72689  Human monocyte che
 11     16      88.9      23     22  AAY72692  Human monocyte che
```

```
                                            12  16  88.9  23  22  AAY72693  Human monocyte che
                                            13  16  88.9  24  22  AAY72682  Human monocyte che
                                            14  16  88.9  26  22  AAY72683  Human monocyte che
                                            15  16  88.9  29  20  AAY12972  TSAR binding doma:
                                            16  16  88.9  29  20  AAR58398  protein sequence o
                                            17  16  88.9  38  15  AAG74890  Human colon cancer
                                            18  16  88.9  49  22  AAY45304  Human secreted pro
                                            19  16  88.9  53  20  AAY48467  Human breast tumou
                                            20  16  88.9  55  20  AAM17762  Peptide #4196 enco
                                            21  16  88.9  59  22  AAM30271  Peptide #4308 enco
                                            22  16  88.9  59  22  AAM05416  Peptide #4098 enco
                                            23  16  88.9  59  22  AAG74441  Human colon cancer
                                            24  16  88.9  60  21  AAB58811  Breast and ovarian
                                            25  16  88.9  66  22  AAW13598  Monocyte chemoattr
                                            26  16  88.9  66  21  AAG55104  Arabidopsis thalia
                                            27  16  88.9  67  18  AAW13599  Monocyte chemoattr
                                            28  16  88.9  68  18  AAW13597  Monocyte chemoattr
                                            29  16  88.9  68  20  AAW95035  MCP-1 analogue [V9
                                            30  16  88.9  68  20  AAW95036  MCP-1 analogue [V9
                                            31  16  88.9  68  20  AAW95037  MCP-1 analogue [V9
                                            32  16  88.9  69  16  AAR87678  des(2-8) MCP-1. H
                                            33  16  88.9  69  18  AAW13596  Monocyte chemoattr
                                            34  16  88.9  69  10  AAP90292  Peptide from human
                                            35  16  88.9  76  13  AAR28660  MCF. Synthetic.
                                            36  16  88.9  76  15  AAR33398  Sense MCP-1. Manm
                                            37  16  88.9  76  16  AAR87680  Monocyte chemotact
                                            38  16  88.9  76  16  AAR87676  (24-Arg) MCP-1. H
                                            39  16  88.9  76  16  AAR87677  (3-Ala) MCP-1. Ho
                                            40  16  88.9  76  17  AAW09374  Monocyte chemotact
                                            41  16  88.9  76  18  AAW11131  Mature human monoc
                                            42  16  88.9  76  19  AAW40175  Macrophage chemoat
                                            43  16  88.9  76  21  AAB12818  Human glioma cell
                                            44  16  88.9  76  21  AAY69030  Amino acid sequenc
                                            45  16  88.9  77  16  AAR86859  Mature MCP-1. Hom
                                            46  16  88.9  77  20  AAY14222  Chemokine hMCP1.
                                            47  16  88.9  78  21  AAB10256  Human adult retina
                                            48  16  88.9  79  22  AAM24442  Human EST encoded
                                            49  16  88.9  90  21  AAG01263  Human secreted pro
                                            50  16  88.9  96  15  AAR51942  Sequence of M. pul
```

ALIGNMENTS

```
RESULT   1
AAY34076

ID     AAY34076 standard; peptide; 7 AA.
XX
AC     AAY34076;
XX
DT     23-NOV-1999  (first entry)
XX
DE     Bacteroides 100 kDa antigen peptide fragment.
XX
KW     Ulcerative colitis; histone; H1-like antigen; porin antigen;
KW     Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW     perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
XX
OS     Bacteroides caccae.
XX
PN     WO9945955-A1.
XX
PD     16-SEP-1999.
XX
PF     12-MAR-1999;  99WO-US05492.
XX
PR     12-MAR-1998;  98US-0041889.
XX
PA     (REGC ) UNIV CALIFORNIA.
XX
PI     Braun J,  Cohavy O;
XX
DR     WPI; 1999-551215/46.
```

RESULT 2

```
AAY34080
ID   AAY34080 standard; peptide; 7 AA.
XX
AC   AAY34080;
XX
DT   23-NOV-1999   (first entry)
XX
DE   Bacteroides 100 kDa antigen peptide fragment.
XX
KW   Ulcerative colitis; histone; Hl-like antigen; porin antigen;
KW   Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW   perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
XX
OS   Bacteroides caccae.
XX
PN   WO9945955-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999;   99WO-US05492.
XX
PR   12-MAR-1998;   98US-0041889.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Braun J, Cohavy O;
XX
DR   WPI; 1999-551215/46.
XX
PT   Use of histone H1, porin or Bacteroides antigens as targets for the
PT   diagnosis, prevention and treatment of ulcerative colitis -
XX
PS   Example 6; Page 78; 134pp; English.
XX
CC   The invention provides a method for the diagnosis, prevention and
CC   treatment of ulcerative colitis (UC) using histone Hl-like antigen, a
CC   porin antigen or a Bacteroides antigen as a target antigen. The novel
CC   method of diagnosing UC in a subject suspected of having inflammatory
CC   bowel disease (IBD) comprises: (1) obtaining a sample from the subject;
CC   (2) contacting the sample with a histone Hl-like antigen, or perinuclear
CC   anti-neutrophil cytoplasmic antibody (pANCA)-reactive fragment, to form
CC   a complex of the histone Hl-like antigen, or the pANCA-reactive
CC   Hl-like antigen, porin antigen and Bacteroides antigen are useful in the
CC   fragment, and antibody to the histone Hl-like antigen; and (3) detecting
CC   the presence or absence of the complex; where the presence of the
CC   complex indicates that the subject has UC. The pANCA-reactive histone
CC   Hl-like antigen, porin antigen and Bacteroides antigen are useful in the
CC   diagnosis, prevention and treatment of UC. The methods can also be used
CC   for identifying agents useful for treating UC. Sequences AAY34072-87
CC   represent peptides derived from the B. caccae 100 kDa antigen (a
CC   microbial pANCA-reactive protein).
XX
SQ   Sequence   7 AA;

Query Match             88.9%;  Score 16;  DB 20;  Length 7;
  Best Local Similarity   60.0%;   Pred. No. 4.3e+05;
  Matches   3;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 LXXYR 5
              | ||
Db          3 lasyr 7
```

RESULT 3

```
AAY34084
ID   AAY34084 standard; peptide; 7 AA.
XX
AC   AAY34084;
XX
DT   23-NOV-1999   (first entry)
XX
DE   Bacteroides 100 kDa antigen peptide fragment.
XX
KW   Ulcerative colitis; histone; Hl-like antigen; porin antigen;
KW   Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW   perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
XX
OS   Bacteroides caccae.
XX
PN   WO9945955-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999;   99WO-US05492.
XX
PR   12-MAR-1998;   98US-0041889.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Braun J, Cohavy O;
XX
DR   WPI; 1999-551215/46.
XX
PT   Use of histone H1, porin or Bacteroides antigens as targets for the
PT   diagnosis, prevention and treatment of ulcerative colitis -
XX
PS   Example 6; Page 78; 134pp; English.
XX
CC   The invention provides a method for the diagnosis, prevention and
CC   treatment of ulcerative colitis (UC) using histone Hl-like antigen, a
CC   porin antigen or a Bacteroides antigen as a target antigen. The novel
CC   method of diagnosing UC in a subject suspected of having inflammatory
CC   bowel disease (IBD) comprises: (1) obtaining a sample from the subject;
CC   (2) contacting the sample with a histone Hl-like antigen, or perinuclear
CC   anti-neutrophil cytoplasmic antibody (pANCA)-reactive fragment, to form
CC   a complex of the histone Hl-like antigen, or the pANCA-reactive
CC   fragment, and antibody to the histone Hl-like antigen; and (3) detecting
CC   the presence or absence of the complex; where the presence of the
CC   complex indicates that the subject has UC. The pANCA-reactive histone
CC   Hl-like antigen, porin antigen and Bacteroides antigen are useful in the
CC   diagnosis, prevention and treatment of UC. The methods can also be used
CC   for identifying agents useful for treating UC. Sequences AAY34072-87
CC   represent peptides derived from the B. caccae 100 kDa antigen (a
CC   microbial pANCA-reactive protein).
XX
SQ   Sequence   7 AA;

Query Match             88.9%;  Score 16;  DB 20;  Length 7;
  Best Local Similarity   60.0%;   Pred. No. 4.3e+05;
  Matches   3;  Conservative   0;  Mismatches   2;  Indels   0;  Gaps   0;

Qy          1 LXXYR 5
              | ||
Db          3 lasyr 7
```

```
OM protein - protein search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:          April 21, 2002, 10:11:15 ; Search time 94.43 Seconds
                                            (without alignments)
                                            1.192 Million cell updates/sec

Title:           US-09-281-760C-1
Perfect score:   18
Sequence:        1 LXXYR 5
Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched:        212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters:    212252
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

     Pred. No. is the number of results predicted by chance to have a
     score greater than or equal to the score of the result being printed,
     and is derived by analysis of the total score distribution.

                          SUMMARIES

Result            %   Query
No.    Score  Match  Length  DB  ID                 Description
-------------------------------------------------------------------------
1      16    88.9    11     6   5177060-30          Patent No. 5177060
2      16    88.9    38     1   US-08-176-500-30    Sequence 56, Appl
3      16    88.9    38     1   US-08-471-052A-56   Sequence 56, Appl
4      16    88.9    38     1   US-08-189-331-56    Sequence 56, Appl
5      16    88.9    38     2   US-08-471-939-56    Sequence 56, Appl
6      16    88.9    38     2   US-08-471-800-56    Sequence 56, Appl
7      16    88.9    38     2   US-08-471-068-56    Sequence 56, Appl
8      16    88.9    69     4   US-08-855-526B-44   Sequence 44, Appl
9      16    88.9    76     1   US-07-956-862A-1    Sequence 1, Appl
10     16    88.9    76     1   US-08-250-958-1     Sequence 1, Appl
11     16    88.9    76     1   US-08-235-659-1     Sequence 1, Appl
12     16    88.9    76     2   US-08-716-188-2     Sequence 2, Appl
13     16    88.9    76     5   US-08-615-332A-5    Sequence 5, Appl
14     16    88.9    76     3   US-08-470-333-5     Sequence 5, Appl
15     16    88.9    78     1   US-08-330-163-12    Sequence 12, Appl
16     16    88.9    78     1   US-08-482-111-12    Sequence 12, Appl
17     16    88.9    78     5   PCT-US95-00605-1    Sequence 1, Appl
18     16    88.9    96     5   PCT-US93-08744-4    Sequence 4, Appl
19     16    88.9    99     1   US-08-127-499A-35   Sequence 35, Appl
20     16    88.9    99     1   US-08-482-847-35    Sequence 35, Appl
21     16    88.9    99     8   US-08-347-492B-8    Sequence 8, Appl
22     16    88.9    99     1   US-08-380-449-19    Sequence 19, Appl
23     16    88.9    99     2   US-08-479-126B-5    Sequence 5, Appl
24     16    88.9    99     2   US-08-421-144A-5    Sequence 5, Appl
25     16    88.9    99     2   US-08-726-830A-5    Sequence 5, Appl
26     16    88.9    99     2   US-08-660-542-19    Sequence 19, Appl
27     16    88.9    99     2   US-08-798-143-8     Sequence 8, Appli
28     16    88.9    99     3   US-07-927-391-24    Sequence 24, Appl
29     16    88.9    99     3   US-08-995-156A-5    Sequence 5, Appl
30     16    88.9    99     3   US-09-044-856A-5    Sequence 5, Appl
31     16    88.9    99     4   US-09-044-855A-5    Sequence 5, Appl
32     16    88.9    99     4   US-08-679-493A-152  Sequence 152, App
33     16    88.9    99     5   PCT-US96-10087-5    Sequence 5, Appl
34     16    88.9    99     6   5212073-2           Patent No. 5212073
35     16    88.9    121    6   US-08-420-235B-23   Sequence 23, Appl
36     16    88.9    121    4   US-08-793-624-23    Sequence 23, Appl
37     16    88.9    121    5   PCT-US95-10194-23   Sequence 23, Appl
38     16    88.9    191    1   US-08-159-784-3     Sequence 3, Appl
39     16    88.9    195    1   US-07-959-284-5     Sequence 5, Appl
40     16    88.9    195    2   US-08-308-736A-5    Sequence 5, Appl
41     16    88.9    195    5   PCT-US92-08258-2    Sequence 2, Appl
42     16    88.9    195    5   PCT-US93-09649A-5   Sequence 5, Appl
43     16    88.9    195    5   PCT-US93-09649-5    Sequence 5, Appl
44     16    88.9    249    5   US-08-591-468-2     Sequence 2, Appl
45     16    88.9    294    5   PCT-US94-06430-2    Sequence 2, Appl
46     16    88.9    312    4   US-09-242-948-2     Sequence 2, Appli
47     16    88.9    317    2   PCT-US94-06430-2    Sequence 2, Appli
48     16    88.9    317    2   US-08-767-096-2     Sequence 2, Appli
49     16    88.9    317    4   US-09-480-203-2     Sequence 2, Appli
50     16    88.9    334    2   US-08-665-647-3     Sequence 3, Appli


                          ALIGNMENTS

RESULT 1
5177060-30
;Patent No. 5177060
;APPLICANT: WEI, EDWARD T.
;TITLE OF INVENTION: ANTI-INFLAMMATORY PEPTIDES AND TREATMENT
;TO INHIBIT VASCULAR LEAKAGE IN INJURED TISSUES
;NUMBER OF SEQUENCES: 30
;CURRENT APPLICATION DATA:
;  APPLICATION NUMBER: US/08/462,578
;  FILING DATE: 09-JAN-1990
;SEQ ID NO:30:
;LENGTH: 11

Query Match      88.9%;  Score 16;  DB 6;  Length 11;
Best Local Similarity  60.0%;  Pred. No. 77;
Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy   1  LXXYR  5
        | ||
Db   1  LATYR  5

RESULT 2
US-08-176-500-56
;Sequence 56, Application US/08176500
;Patent No. 5498538
;GENERAL INFORMATION:
;APPLICANT:
;  APPLICANT: Kay, B. K.
;  APPLICANT: Fowlkes, D. M.
;TITLE OF INVENTION: Totally Synthetic Affinity Reagents
;NUMBER OF SEQUENCES: 141
;CORRESPONDENCE ADDRESS:
;  ADDRESSEE: Pennie & Edmonds
;  STREET: 1155 Avenue of the Americas
;  CITY: New York
;  STATE: New York
;  COUNTRY: U.S.A.
;  ZIP: 10036-2711
;COMPUTER READABLE FORM:
;  MEDIUM TYPE: Floppy disk
;  COMPUTER: IBM PC compatible
;  OPERATING SYSTEM: PC-DOS/MS-DOS
;  SOFTWARE: PatentIn Release #1.0, Version #1.25
```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/176,500
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/013,416
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-143
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
US-08-176-500-56
MOLECULE TYPE: peptide
TOPOLOGY: unknown
STRANDEDNESS: Single
TYPE: amino acid
LENGTH: 38 amino acids
SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 56:

Query Match   88.9%; Score 16; DB 1; Length 38;
Best Local Similarity   60.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 LXXYR 5
        | ||
Db 26 LTSYR 30

RESULT 3
US-08-471-052A-56
Sequence 56, Application US/08471052A
Patent No. 5625033
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 166
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/471,052A
FILING DATE: 06-JUNE-1995
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-179
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: unknown

MOLECULE TYPE: peptide
US-08-471-052A-56

Query Match   88.9%; Score 16; DB 1; Length 38;
Best Local Similarity   60.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 LXXYR 5
        | ||
Db 26 LTSYR 30

RESULT 4
US-08-189-331-56
Sequence 56, Application US/08189331
Patent No. 5747334
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents
NUMBER OF SEQUENCES: 186
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/189,331
FILING DATE: Concurrently herewith
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-155
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212 790-9090
TELEFAX: 212 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 56:
SEQUENCE CHARACTERISTICS:
LENGTH: 38 amino acids
TYPE: amino acid
STRANDEDNESS: Single
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-08-189-331-56

Query Match   88.9%; Score 16; DB 1; Length 38;
Best Local Similarity   60.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 LXXYR 5
        | ||
Db 26 LTSYR 30

RESULT 5
US-08-471-939-56
Sequence 56, Application US/08471939
Patent No. 5844076
GENERAL INFORMATION:
APPLICANT: Kay, B. K.
APPLICANT: Fowlkes, D. M.
TITLE OF INVENTION: Totally Synthetic Affinity Reagents

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on:　　　　April 21, 2002, 10:17:53 ; Search time 102.84 Seconds
　　　　　　　　　　(without alignments)
　　　　　　　　　　3.704 Million cell updates/sec

Title:　　　　　US-09-281-760C-2
Perfect score:　18
Sequence:　　　　1 YRXXL 5

Scoring table:　BLOSUM62
　　　　　　　　　Gapop 10.0 , Gapext 0.5

Searched:　　　　219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:　219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
　　　　　　　　　Maximum Match 100%
　　　　　　　　　Listing first 50 summaries

Database :　　　PIR_68:*
　　　　　　　1:　pir1:*
　　　　　　　2:　pir2:*
　　　　　　　3:　pir3:*
　　　　　　　4:　pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB	ID	Description
1	16	88.9	56	2	T39209	very hypothetical
2	16	88.9	68	2	S22327	gene C protein - p
3	16	88.9	75	2	H59096	hypothetical protein
4	16	88.9	81	2	S01059	hypothetical prote
5	16	88.9	83	2	A49355	cytochrome-c oxida
6	16	88.9	83	4	T14380	probable pollen co
7	16	88.9	101	4	S59321	hypothetical prote
8	16	88.9	102	4	C96595	hypothetical prote
9	16	88.9	102	2	A84266	hypothetical prote
10	16	88.9	112	2	C75298	hypothetical prote
11	16	88.9	114	2	F85843	unknown protein en
12	16	88.9	114	2	B85692	unknown protein en
13	16	88.9	114	2	B85681	hypothetical prote
14	16	88.9	114	2	B85630	hypothetical prote
15	16	88.9	118	2	T05812	hypothetical prote
16	16	88.9	119	2	MSPG	motilin precursor
17	16	88.9	124	2	B48983	NADH dehydrogenase
18	16	88.9	130	2	H69699	ribosomal protein
19	16	88.9	130	2	T47309	hypothetical protein
20	16	88.9	131	2	H69062	molybdenum transpo
21	16	88.9	131	4	A84491	hypothetical prote
22	16	88.9	132	2	S77266	glycine cleavage s
23	16	88.9	133	4	S24315	motilin precursor
24	16	88.9	139	4	T03486	conserved hypothet
25	16	88.9	151	2	S69472	hypothetical prote
26	16	88.9	152	2	T44544	conserved hypothet
27	16	88.9	156	2	E84230	hypothetical prote
28	16	88.9	159	2	E83346	probable sigma-70
29	16	88.9	160	2	T04633	hypothetical prote
30	16	88.9	175	2	D86471	hypothetical prote
31	16	88.9	176	2	T47361	hypothetical prote
32	16	88.9	182	2	T49540	hypothetical prote
33	16	88.9	184	2	A83722	hypothetical prote
34	16	88.9	185	2	A28677	aminoglycoside N6'
35	16	88.9	188	2	D75499	probable acetyltra
36	16	88.9	189	2	T05505	lds-4 protein - ba
37	16	88.9	189	2	D83008	outer membrane 1p
38	16	88.9	189	2	C71427	hypothetical prote
39	16	88.9	190	2	T28093	ribosomal protein
40	16	88.9	194	2	T11928	conserved hypothet
41	16	88.9	195	2	A75295	HetI protein Xf193
42	16	88.9	197	2	C82619	repressor protein
43	16	88.9	198	2	JQ1052	conjugal transfer
44	16	88.9	199	2	D82865	conjugal transfer
45	16	88.9	200	2	D83116	50S ribosomal prot
46	16	88.9	201	2	F82604	conjugal transfer
47	16	88.9	202	2	G71295	conserved hypothet
48	16	88.9	203	2	T18653	hypothetical prote
49	16	88.9	206	2	E75345	peptide methionine
50	16	88.9	208	2	T30859	fertility inhibiti

ALIGNMENTS

RESULT　1
T39209
very hypothetical protein SPAC9E9.17c - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T39209
R;McDougall, R.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
submitted to the EMBL Data Library, September 1997
A;Reference number: Z21836
A;Accession: T39209
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-56 <MCD>
A;Cross-references: EMBL:Z99262; PIDN:CAB58229.1; GSPDB:GN00066; SPDB:SPAC9E9.17c
A;Experimental source: strain 972h-; cosmid c9E9
A;Genetics:
C;Genetics:
A;Gene: SPDB:SPAC9E9.17c
A;Map position: 1
A;Introns: 2/2

　Query Match　　88.9%;　Score 16;　DB 2;　Length 56;
　Best Local Similarity　60.0%;
　Matches　　3;　Conservative　0;　Mismatches　2;　Indels　0;　Gaps　0;

Qy　　1 YRXXL 5
　　　　| |　|
Db　37 YRTTL 41

RESULT　2
S22327
gene C protein - phage alpha-3
C;Species: phage alpha-3
C;Date: 10-Mar-1994 #sequence_revision 10-Mar-1994 #text_change 20-Sep-1999
C;Accession: S22327
R;Kodaira, K.I.; Nakano, K.; Okada, S.; Taketo, A.
Biochim. Biophys. Acta 1130, 277-288, 1992
A;Title: Nucleotide sequence of the genome of the bacteriophage alpha3: interrelation
A;Reference number: S22324; MUID:92223109
A;Accession: S22327
A;Molecule type: DNA
A;Residues: 1-68 <KOD>
A;Cross-references: EMBL:X60322; NID:g14775; PIDN:CAA42877.1; PID:g14779
C;Genetics:
A;Gene: C
C;Superfamily: phage phi-X174 gene C protein

Query Match   88.9%; Score 16; DB 2; Length 68;
Best Local Similarity 60.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy   1 YRXXL 5
        | | |
Db   11 YRATL 15

RESULT 3
H59096
hypothetical protein pXO1-48 - Bacillus anthracis virulence plasmid pXO1
C;Species: Bacillus anthracis
C;Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 11-May-2000
C;Accession: H59096
R;Okinaka, R.T.; Cloud, K.; Hampton, O.; Hoffmaster, A.R.; Hill, K.K.; Keim, P.; Koehler
J;Bacteriol. 181, 6509-6515, 1999
A;Title: Sequence and organization of pXO1, the large Bacillus anthracis plasmid harbori
A;Reference number: H59096
A;Accession: H59096
A;Molecule type: DNA
A;Status: preliminary
A;Residues: 1-75 <OKI>
A;Cross-references: GB;AF065404; NID:9489421 6; PIDN:AAD32352.1; PID:9489426 4
A;Experimental source: strain Sterne
C;Genetics:
A;Gene: pXO1-48
A;Genome: plasmid

Qy   1 YRXXL 5
        | | |
Db   70 YRTTL 74

Query Match   88.9%; Score 16; DB 2; Length 75;
Best Local Similarity 60.0%; Pred. No. 5.4e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 4
S01059
hypothetical protein 2 - rice
C;Species: Oryza sativa (rice)
C;Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
C;Accession: S01059
R;Kikuchi, S.; Takaiwa, F.; Oono, K.
J;Mol. Gen. Genet. 210, 373-380, 1987
A;Title: Variable copy number DNA sequences in rice.
A;Reference number: S01057; MUID:88121710
A;Accession: S01059
A;Status: translation not shown
A;Molecule type: DNA
A;Residues: 1-81 <KIK>
A;Cross-references: EMBL;X06611; NID:g20380; PIDN:CAA29826.1; PID:g20383

Qy   1 YRXXL 5
        | | |
Db   3 YRSSL 7

Query Match   88.9%; Score 16; DB 2; Length 81;
Best Local Similarity 60.0%; Pred. No. 5.9e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 5
A49355
cytochrome-c oxidase (EC 1.9.3.1) chain VIIa precursor, isoform L - bovine
N;Alternate names: cytochrome-c oxidase chain VIIa, hepatic
C;Species: Bos primigenius taurus (cattle)
C;Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 11-Jun-1999

C;Accession: A49355; B41034; S07891; C29968
R;Seelan, R.S.; Grossman, L.I.
G;Genomics 18, 527-536, 1993
A;Title: Structural organization and evolution of the liver isoform gene for bovine c
A;Reference number: A49355
A;Accession: A49355
A;Molecule type: DNA
A;Residues: 1-83 <SE3>
A;Cross-references: GB;L09603; NID:g451884; PIDN:AAA18218.1; PID:g488158
R;Seelan, R.S.; Grossman, L.I.
J;Biol. Chem. 266, 19752-19757, 1991
A;Title: Cytochrome c oxidase subunit VIIa isoforms. Characterization and expression
A;Accession: B41034
A;Reference number: B41034
A;Molecule type: mRNA
A;Residues: 1-83 <SEE>
A;Note: this corrects the sequence reported with accession number S07891
R;Seelan, R.S.; Scheuner, D.; Lomax, M.I.; Grossman, L.I.
N;Nucleic Acids Res. 17, 6410, 1989
A;Title: Nucleotide sequence of a cDNA for bovine cytochrome c oxidase subunit VIIa.
A;Reference number: S07891; MUID:89366688
A;Accession: S07891
A;Molecule type: mRNA
A;Residues: 1-52, /R', 54-83 <SE2>
A;Cross-references: EMBL;X15235; NID:g273; PIDN:CAA33313.1; PID:g274
R;Yanamura, W.; Zhang Y.Z.; Takamiya, S.; Capaldi, R.A.
J;Biochemistry 27, 4909-4914, 1988
A;Title: Tissue-specific differences between heart and liver cytochrome c oxidase.
A;Reference number: A90531; MUID:89000697
A;Accession: C29968
A;Molecule type: protein
A;Residues: 24-51, 'NI', 54-56, 'V', 58, 'M', 60 <YAN>
A;Experimental source: liver
C;Genetics:
A;Gene: COX7AL
A;Introns: 6/3; 36/3; 65/1
C;Superfamily: mammalian cytochrome-c oxidase chain VIIa
C;Keywords: liver; mitochondrion; oxidoreductase
F;1-23/Domain: transit peptide (mitochondrion) #status predicted <TNP>
F;24-83/Product: cytochrome-c oxidase chain VIIa #status predicted <MAT>

Qy   1 YRXXL 5
        | | |
Db   55 YRATL 59

Query Match   88.9%; Score 16; DB 2; Length 83;
Best Local Similarity 60.0%; Pred. No. 6e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 6
T14380
probable pollen coat protein 1 - turnip
C;Species: Brassica rapa (turnip)
C;Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 21-Jul-2000
C;Accession: T14380
R;Toriyama, K.; Hanaoka, K.; Okada, T.; Watanabe, M.
FEBS Lett. 424, 234-238, 1998
A;Title: Molecular cloning of a cDNA encoding a pollen extracellular protein as a pot
A;Reference number: T14380
A;Accession: T14380
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: mRNA
A;Residues: 1-83 <TOR>
A;Cross-references: EMBL;AB010435; NID:g3062794; PIDN:BAA25682.1; PID:g3062795
C;Genetics:
A;Experimental source: anther
A;Gene: PEC-3

Query Match   88.9%; Score 16; DB 2; Length 83;

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:37:20 ; Search time 85.5 Seconds
                 (without alignments)
                 2.144 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5

Scoring table: BLOSUM62
Scoring: Gapop 10.0 , Gapext 0.5

Database :   SwissProt_39:*

Searched: 100059 seqs, 3664827 residues

Total number of hits satisfying chosen parameters: 100059

Optimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match Length	DB	ID		Description
1	16	88.9	68	1	VGC_BPAL3	P31279	bacteriopha
2	16	88.9	68	1	VGC_BPPHK	Q38037	bacteriopha
3	16	88.9	83	1	COXJ_BOVIN	P13184	bos taurus
4	16	88.9	92	1	MOTI_HORSE	O46617	equus cabal
5	16	88.9	111	1	VP34_BPAPS	Q91r4	bacteriopha
6	16	88.9	115	1	MOTI_MACMU	O18811	macaca mula
7	16	88.9	116	1	MOTI_FELCA	Q9xse2	felis silve
8	16	88.9	119	1	MOTI_PIG	Q01307	sus scrofa
9	16	88.9	124	1	NU5C_PEA	Q32905	pisum sativ
10	16	88.9	129	1	RS9_BACSU	P21470	bacillus su
11	16	88.9	130	1	UCR6_SCHMA	P21474	schistosoma
12	16	88.9	132	1	GCSH_SYNY3	P73360	synechocyst
13	16	88.9	133	1	MOTI_RABIT	P27114	oryctolagus
14	16	88.9	183	1	RL6_MYCGA	O52347	mycoplasma
15	16	88.9	185	1	AAC6_CITDI	P10051	citrobacter
16	16	88.9	188	1	VATE_THETH	P74901	thermus.aqu
17	16	88.9	194	1	RM06_PROWI	P46748	prototheca
18	16	88.9	198	1	OPNR_AGRRH	P27872	agrobacteri
19	16	88.9	206	1	MSRA_DEIRA	Q9rtb6	deinococcus
20	16	88.9	208	1	RL13_CHLSW	O48513	chlamydomon
21	16	88.9	210	1	SGCE_ECOLI	P39362	escherichia
22	16	88.9	213	1	RM06_SCHPO	Q09865	schizosacch
23	16	88.9	214	1	VTI1_SCHPO	P78768	schizosacch
24	16	88.9	229	1	Y454_TREPA	O83467	treponema p
25	16	88.9	230	1	CUTI_FUSSC	Q99174	fusarium so
26	16	88.9	230	1	CUTI_FUSSO	P00590	fusarium so
27	16	88.9	234	1	CITB_KLEPN	P52688	klebsiella
28	16	88.9	236	1	HO_PORPU	P51271	porphyra pu
29	16	88.9	239	1	RS3_SYNY3	P73314	synechocyst
30	16	88.9	245	1	END3_MYCLE	Q9cb92	mycobacteri
31	16	88.9	248	1	Y4EK_RHISN	P55434	rhizobium s
32	16	88.9	263	1	YEMT_HUMAN	Q95050	homo sapien
33	16	88.9	264	1	NNMT_MOUSE	O55239	mus musculu
34	16	88.9	264	1	TEMT_MOUSE	P40936	mus musculu
35	16	88.9	269	1	MIND_GUITH	O78436	guillardia
36	16	88.9	271	1	ELH1_APLCA	P01362	aplysia cal
37	16	88.9	276	1	SSUC_BACSU	P40401	bacillus su
38	16	88.9	279	1	YEB8_YEAST	P39995	saccharomyc
39	16	88.9	293	1	VG11_BPB03	Q37892	bacteriopha
40	16	88.9	296	1	KHSE_LACLA	Q9cgd7	lactococcus
41	16	88.9	296	1	KHSE_LACLC	P52991	lactococcus
42	16	88.9	298	1	YSMK_CAEEL	Q19408	caenorhabdi
43	16	88.9	300	1	PPRA_DEIRA	O32504	deinococcus
44	16	88.9	310	1	S6PD_MALDO	P28475	malus domes
45	16	88.9	313	1	MPC1_ALCEU	P17295	alcaligenes
46	16	88.9	315	1	SR2A_PHYPO	P09352	physarum po
47	16	88.9	316	1	THER_BACTH	P00800	bacillus th
48	16	88.9	322	1	PIV_MORLA	P19257	moraxella l
49	16	88.9	324	1	SECF_RHOSH	O33568	rhodobacter
50	16	88.9	325	1	HLGB_STAAU	Q07226	staphylococ

ALIGNMENTS

RESULT 1
VGC_BPAL3
ID   VGC_BPAL3   STANDARD;   PRT;   68 AA.
AC   P31279;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-JUN-1994 (Rel. 29, Last annotation update)
DE   PROTEIN.
GN   C.
OS   Bacteriophage alpha-3.
OC   Viruses; ssDNA viruses; Microviridae; Microvirus.
OX   NCBI_TaxID=10849;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92223109; PubMed=1532908;
RA   Kodaira K.-I., Nakano K., Okada S., Taketo A.;
RT   "Nucleotide sequence of the genome of the bacteriophage alpha 3:
RT   interrelationship of the genome structure and the gene products with
RT   those of the phages. phi X174, G4 and phi K.";
RL   Biochim. Biophys. Acta 1130:277-288(1992).
CC   -!- FUNCTION: C PROTEIN IS ONE OF THE PROTEINS INVOLVED IN THE
CC   PRODUCTION AND PACKAGING OF VIRAL SINGLE-STRANDED DNA.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; X60322; CAA42877.1; -.
DR   PIR; S22327; S22327.
SQ   SEQUENCE   68 AA;   8295 MW;   7F8OD3EA2FBB4D4F CRC64;

  Query Match   88.9%;   Score 16;   DB 1;   Length 68;
  Best Local Similarity 60.0%;   Pred. No. 1.7e+02;
  Matches   3;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 YRXXL 5
         || |
Db   11 YRAYL 15

RESULT 2
VGC_BPPHK
ID   VGC_BPPHK   STANDARD;   PRT;   68 AA.
AC   Q38037;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)

DT   01-NOV-1997 (Rel. 35, Last annotation update)
DE   C PROTEIN.
GN   C.
OS   Bacteriophage phi-K.
OC   Viruses; ssDNA viruses; Microviridae; Microvirus.
OX   NCBI_TaxID=10848;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kodaira K.-I., Tadokoro H., Suzuki K., Oki M., Manto S., Taketo A.;
RL   Submitted (JUL-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: C PROTEIN IS ONE OF THE PROTEINS INVOLVED IN THE
CC   PRODUCTION AND PACKAGING OF VIRAL SINGLE-STRANDED DNA.
CC   --

DR   EMBL; X60323; CAA42886.1; -.
SQ   SEQUENCE 68 AA; 8323 MW; 7F80D3EA388B4D4F CRC64;

Query Match
Best Local Similarity   88.9%;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;
                                 Score 16; DB 1; Length 68; Pred. No. 1.7e+02;

Qy   1 YRXXL 5
            | | |
Db   11 YRATL 15

RESULT 3

ID   COXJ_BOVIN STANDARD; PRT; 83 AA.
AC   P13184;
DT   01-JAN-1990 (Rel. 13, Created)
DT   01-DEC-1992 (Rel. 24, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIIA-LIVER/HEART, MITOCHONDRIAL
DE   PRECURSOR (EC 1.9.3.1) (CYTOCHROME C OXIDASE SUBUNIT VIIA-L).
GN   COX7A2 OR COX7AL.
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
OC   Bovidae; Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89366688; PubMed=2549516;
RA   Seelan R.S., Scheuner D., Lomax M.I., Grossman L.I.;
RT   "Nucleotide sequence of a cDNA for bovine cytochrome c oxidase
RT   subunit VIIa.";
RL   Nucleic Acids Res. 17:6410-6410(1989).
RN   [2]
RP   REVISIONS, SEQUENCE FROM N.A.
RX   MEDLINE=92011781; PubMed=1717471;
RA   Seelan R.S., Grossman L.I.;
RT   "Cytochrome c oxidase subunit VIIa isoforms. Characterization and
RT   expression of bovine cDNAs.";
RL   J. Biol. Chem. 266:19752-19757(1991).
RN   [3]
RP   SEQUENCE OF 24-60.
RX   MEDLINE=89006097; PubMed=2844245;
RA   Yanamura W., Zhang Y.-Z., Takamiya S., Capaldi R.A.;
RT   "Tissue-specific differences between heart and liver cytochrome c
RT   oxidase.";
RL   Biochemistry 27:4909-4914(1988).
CC   -!- FUNCTION: THIS PROTEIN IS ONE OF THE NUCLEAR-CODED POLYPEPTIDE
CC   CHAINS OF CYTOCHROME C OXIDASE, THE TERMINAL OXIDASE IN
CC   MITOCHONDRIAL ELECTRON TRANSPORT.
CC   -!- CATALYTIC ACTIVITY: 4 FERROCYTOCHROME C + O(2) = 2 H(2)O +
CC   4 FERRICYTOCHROME C.
CC   -!- SUBCELLULAR LOCATION: MITOCHONDRIAL INNER MEMBRANE.
CC   -!- SIMILARITY: BELONGS TO THE CYTOCHROME C OXIDASE VIIA FAMILY.
CC   --

DR   EMBL; L09603; AAA18218.1; -.
DR   EMBL; X15225; CAA33313.1; ALT_SEQ.
DR   PIR; B41034; B41034.
DR   PIR; C29960; C29960.
DR   Interpro; IPR003177; COX7a.
DR   Pfam; PF02238; COX7a.
KW   Oxidoreductase; Inner membrane; Mitochondrion; Transit peptide.
FT   TRANSIT   1   23   MITOCHONDRION; Transit peptide.
FT   CHAIN   24   83   CYTOCHROME C OXIDASE POLYPEPTIDE
FT   VIIA-LIVER/HEART.
FT   CONFLICT   52   53   AL -> NI (IN REF. 3).
FT   CONFLICT   57   57   A -> V (IN REF. 3).
FT   CONFLICT   59   59   L -> M (IN REF. 3).
SQ   SEQUENCE 83 AA; 9305 MW; 254D7610F2EF106F CRC64;

Query Match
Best Local Similarity   88.9%;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;
                                 Score 16; DB 1; Length 83; Pred. No. 2.1e+02;

Qy   1 YRXXL 5
            | | |
Db   55 YRATL 59

RESULT 4

ID   MOTI_HORSE STANDARD; PRT; 92 AA.
AC   O46617;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   MOTILIN PRECURSOR [CONTAINS: MOTILIN; MOTILIN ASSOCIATED PEPTIDE
DE   (MAP)] (FRAGMENT).
GN   MLN.
OS   Equus caballus (Horse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX   NCBI_TaxID=9796;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Duodenal mucosa;
RA   Huang Z., De Clercq P., Depoortere I., Peeters T.L.;
RT   "cDNA encoding motilin precursor from horse duodenum mucosa.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE REGULATION OF
CC   INTERDIGESTIVE GASTROINTESTINAL MOTILITY AND INDIRECTLY CAUSES
CC   RHYTHMIC CONTRACTION OF DUODENAL AND COLONIC SMOOTH MUSCLE (BY
CC   SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: SECRETED.
CC   -!- SIMILARITY: BELONGS TO THE MOTILIN FAMILY.
CC   --

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:35:42 ; Search time 191.29 Seconds
                         (without alignments)
                         3.823 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   SPTREMBL_17:*
               1: sp_archea:*
               2: sp_bacteria:*
               3: sp_fungi:*
               4: sp_human:*
               5: sp_invertebrate:*
               6: sp_mammal:*
               7: sp_mhc:*
               8: sp_organelle:*
               9: sp_phage:*
              10: sp_plant:*
              11: sp_rodent:*
              12: sp_virus:*
              13: sp_vertebrate:*
              14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	88.9	44	8	P92073	euhadra her
2	16	88.9	56	2	Q9UT14	schizosacch
3	16	88.9	63	2	P94133	acinetobact
4	16	88.9	63	5	Q9TV14	oryctolagus
5	16	88.9	66	5	Q9NM53	leishmania
6	16	88.9	73	2	Q9F6T0	helicobacte
7	16	88.9	73	2	Q9EUA9	helicobacte
8	16	88.9	73	2	Q9EU25	helicobacte
9	16	88.9	75	2	Q9X318	bacillus an
10	16	88.9	81	10	Q40745	oryza sativ
11	16	88.9	83	10	O64433	brassica ra
12	16	88.9	83	10	Q9MB53	brassica na
13	16	88.9	83	10	Q9MB52	brassica ju
14	16	88.9	83	10	Q9LD24	brassica ca
15	16	88.9	90	12	O41642	human immun
16	16	88.9	93	8	Q9UVI0	gibberella
17	16	88.9	93	8	Q9MTL9	oenothera h
18	18	88.9	101	3	O07236	saccharomyc
19	16	88.9	102	1	Q9HQM0	halobacteri
20	16	88.9	102	10	Q9C8A2	arabidopsis
21	16	88.9	102	11	Q9QVZ1	rattus sp.
22	16	88.9	108	5	O61383	papilio gla
23	16	88.9	112	2	Q9RS90	deinococcus
24	16	88.9	113	13	Q9I8A6	meleagris g
25	16	88.9	114	2	P96462	salmonella
26	16	88.9	115	9	Q9MCS8	bacteriopha
27	16	88.9	117	2	Q9Z5T2	zymomonas m
28	16	88.9	117	8	O65700	arabidopsis
29	16	88.9	118	11	Q9D8E1	mus musculu
30	16	88.9	122	5	Q9GU26	sanderia ma
31	16	88.9	126	2	Q9Z5Y0	frankia sp.
32	16	88.9	128	5	Q9NH50	peneus mon
33	16	88.9	130	10	Q9FGB4	arabidopsis
34	16	88.9	130	2	Q9RJF7	streptomyce
35	16	88.9	131	10	Q9M1M5	arabidopsis
36	16	88.9	131	1	O27514	methanobact
37	16	88.9	131	10	Q9SHW1	arabidopsis
38	16	88.9	139	2	O68053	rhodobacter
39	16	88.9	140	12	O11312	molluscum c
40	16	88.9	143	11	Q9D9X3	mus musculu
41	16	88.9	152	2	Q9S575	molluscum c
42	16	88.9	156	1	Q9HRE9	halobacteri
43	16	88.9	159	2	Q9I193	pseudomonas
44	16	88.9	160	10	O49380	arabidopsis
45	16	88.9	165	10	Q9FGB2	arabidopsis
46	16	88.9	166	12	O66019	crimean-con
47	16	88.9	167	10	Q9LT29	arabidopsis
48	16	88.9	169	5	O61605	drosophila
49	16	88.9	171	11	Q9CUN9	mus musculu
50	16	88.9	172	2	P97025	brevibacter

ALIGNMENTS

RESULT 1
P92073

ID	P92073	PRELIMINARY;	PRT;	44 AA.

AC   P92073;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ATPASE SUBUNIT 6 (FRAGMENT).
OS   Euhadra herklotsi.
OG   Mitochondrion.
OC   Eukaryota; Metazoa; Mollusca; Gastropoda; Pulmonata; Stylommatophora;
OC   Helicacea; Bradybaenidae; Euhadra.
OX   NCBI_TaxID=58912;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=HEPATOPANCREAS;
RA   Yamazaki N., Ueshima R., Terrett J.A., Yokobori S., Kaifu M.,
RA   Segawa R., Kobayashi T., Numachi K., Ueda T., Nishikawa K.,
RA   Watanabe K., Thomas R.H.;
RL   Submitted (MAY-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z71697; CAA96374.1; -.
DR   InterPro; IPR000568; ATP-synt_A.
DR   Pfam; PF00119; ATP-synt_A; 1.
KW   Mitochondrion.
FT   NON_TER   1
FT   NON_TER   1
SQ   SEQUENCE   44 AA;   5279 MW;   6EFB539307F2DF89 CRC64;

Query Match   88.9%;
Best Local Similarity   60.0%;
Matches   3;   Conservative

    Score 16;   DB 8;   Length 44;
    Pred. No. 9.1e+02;
    0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 YRXXL 5
         Y || |
Db   39 YRSSL 43

RESULT 2
Q9TV14
ID   Q9TV14   PRELIMINARY;   PRT;   56 AA.
AC   Q9TV14;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   VERY HYPOTHETICAL PROTEIN.
GN   SPAC9E9.17C.
OS   Schizosaccharomyces pombe (Fission yeast).
OC   Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC   Schizosaccharomycetales; Schizosaccharomycetaceae;
OC   Schizosaccharomyces.
OX   NCBI_TaxID=4896;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   McDougall R.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Barrell B.G., Rajandream M.A., Wood V.;
RL   Submitted (SEP-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; Z99262; CAB58229.1; -.
SQ   SEQUENCE   56 AA;   6413 MW; 48B23E9B5F2F2676 CRC64;

Query Match   88.9%; Score 16; DB 3; Length 56;
Best Local Similarity   60.0%; Pred. No. 1.2e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 YRXXL 5
        || |
Db   37 YRTTL 41

RESULT 3
P94133
ID   P94133   PRELIMINARY;   PRT;   63 AA.
AC   P94133;
DT   01-MAY-1997 (TrEMBLrel. 03, Created)
DT   01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   HYPOTHETICAL 7.1 KDA PROTEIN (FRAGMENT).
OS   Acinetobacter sp. ADP1.
OC   Bacteria; Proteobacteria; gamma subdivision; Moraxellaceae;
OC   Acinetobacter.
OX   NCBI_TaxID=62977;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=ADP1;
RX   MEDLINE=97440147; PubMed=9294455;
RA   Williams P.A., Shaw L.E.;
RT   "muck, a gene in Acinetobacter calcoaceticus ADP1 (BD413), encodes the
RT   ability to grow on exogenous cis,cis-muconate as the sole carbon
RT   source.";
RL   J. Bacteriol. 179:5935-5942(1997).
DR   EMBL; U87258; AAC27119.1; -.
KW   Hypothetical protein.
FT   NON_TER   1   1
SQ   SEQUENCE   63 AA;   7128 MW; 49B15779381B69B2 CRC64;

Query Match   88.9%; Score 16; DB 2; Length 63;
Best Local Similarity   60.0%; Pred. No. 1.3e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 YRXXL 5
        || |
Db   22 YRTAL 26

RESULT 4
Q9TV14
ID   Q9TV14   PRELIMINARY;   PRT;   63 AA.
AC   Q9TV14;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   POTASSIUM CHANNEL KV4.2 (FRAGMENT).
OS   Oryctolagus cuniculus (Rabbit).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX   NCBI_TaxID=9986;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Shi H., Wang Z.;
RT   "Rabbit potassium channel Kv4.2 mRNA, partial cds.";
RL   Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF085171; AAD49438.1; -.
DR   InterPro; IPR003131; K_tetra.
DR   Pfam; PF02214; K_tetra; 1.
KW   Ionic channel.
FT   NON_TER   63   63
SQ   SEQUENCE   63 AA;   7091 MW; B25335CEBE689CD7 CRC64;

Query Match   88.9%; Score 16; DB 6; Length 63;
Best Local Similarity   60.0%; Pred. No. 1.3e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 YRXXL 5
        || |
Db   22 YRSTL 26

RESULT 5
Q9NM53
ID   Q9NM53   PRELIMINARY;   PRT;   66 AA.
AC   Q9NM53;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   HYPOTHETICAL 7.9 KDA PROTEIN (FRAGMENT).
GN   LM26.239.
OS   Leishmania major.
OC   Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
OX   NCBI_TaxID=5664;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=FRIEDLIN;
RA   Murphy L., Quail M., Harris D., Rajandream M., Ivens A., Barrell B.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL160493; CAB97857.1; -.
KW   Hypothetical protein.
FT   NON_TER   66   66
SQ   SEQUENCE   66 AA;   7876 MW; A3CA0216D54D3C73 CRC64;

Query Match   88.9%; Score 16; DB 5; Length 66;
Best Local Similarity   60.0%; Pred. No. 1.4e+03;
Matches   3; Conservative   0; Mismatches   2; Indels   0; Gaps   0;

Qy   1 YRXXL 5
        || |
Db   60 YRSSL 64

RESULT 6
Q9F6T0
ID   Q9F6T0   PRELIMINARY;   PRT;   73 AA.
AC   Q9F6T0;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:14:05 ; Search time 158.02 Seconds
(without alignments)
2.344 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : A_Geneseq_1101:*

1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	16	88.9	9	18	AAW12322	HiNOS immunogenic
2	16	88.9	10	18	AAW23311	Human inducible ni
3	16	88.9	10	18	AAW12319	Radiolabelled HiNO
4	16	88.9	10	18	AAW12320	C-terminally HiNO
5	16	88.9	10	18	AAW12321	C-terminally subst
6	16	88.9	10	18	AAW11553	Peptide NO54. Syn
7	16	88.9	13	22	AAG98760	Human cell death p
8	16	88.9	30	22	AAG98758	Human cell death p
9	16	88.9	34	22	AAM34167	Peptide #8204 enco
10	16	88.9	35	20	AAY26935	Varicella zoster v
11	16	88.9	35	22	AAM18882	Peptide #5316 enco
12	16	88.9	35	22	AAB73931	Varicella zoster v
13	16	88.9	40	20	AAW96838	Nucleic acid bindi
14	16	88.9	42	22	AAM14043	Peptide #477 encod
15	16	88.9	42	22	AAM26449	Peptide #486 encod
16	16	88.9	42	22	AAM01783	Peptide #465 encod
17	16	88.9	52	20	AAW96852	Nucleic acid bindi
18	16	88.9	62	22	AAM15759	Peptide #2193 enco
19	16	88.9	62	22	AAM28268	Peptide #2305 enco
20	16	88.9	62	22	AAM03493	Peptide #2175 enco
21	16	88.9	65	22	AAM34343	Peptide #8380 enco
22	16	88.9	69	22	AAG73544	Human colon cancer
23	16	88.9	70	20	AAY19484	Amino acid sequenc
24	16	88.9	101	22	AAM33274	Peptide #7311 enco
25	16	88.9	103	21	AAB43078	Human ORFX ORF2842
26	16	88.9	104	22	AAM31485	Peptide #5522 enco
27	16	88.9	105	21	AAB42644	Human ORFX ORF2408
28	16	88.9	117	20	AAY12157	Human 5' EST secre
29	16	88.9	119	22	AAP91380	Porcine prepromoti
30	16	88.9	119	10	AAB62651	Motilin receptor,
31	16	88.9	119	22	AAB20102	Pig motilin. Sus
32	16	88.9	121	21	AAB38583	Human secreted pro
33	16	88.9	122	22	AAG71349	Human secreted pro
34	16	88.9	124	22	AAG72506	Human OR-like poly
35	16	88.9	125	21	AAB18103	Pinus radiata pine
36	16	88.9	125	21	AAG15006	Arabidopsis thalia
37	16	88.9	125	22	AAG73819	Human colon cancer
38	16	88.9	128	22	AAB55819	Human colon cancer
39	16	88.9	128	22	AAB65824	Murine TANGO 253 C
40	16	88.9	131	21	AAG14926	Arabidopsis thalia
41	16	88.9	131	21	AAG14925	Arabidopsis thalia
42	16	88.9	133	21	AAG14924	Arabidopsis thalia
43	16	88.9	136	21	AAB11134	Bordetella pertuss
44	16	88.9	139	21	AAG15005	Arabidopsis thalia
45	16	88.9	142	21	AAB41090	Human ORFX ORF854
46	16	88.9	146	21	AAG54025	Arabidopsis thalia
47	16	88.9	155	22	AAG82284	S. epidermidis ope
48	16	88.9	156	21	AAG23444	Arabidopsis thalia
49	16	88.9	158	21	AAG54019	Arabidopsis thalia
50	16	88.9	163	21	AAB25318	Eucalyptus grandis

ALIGNMENTS

RESULT 1
AAW12322

ID AAW12322 standard; Peptide; 9 AA.
XX
AC AAW12322;
XX
DT 20-OCT-1997 (first entry)
XX
DE HiNOS immunogenic peptide without C-terminal leucine.
XX
KW C-terminal; epitope; human; inducible nitric oxide synthase; hiNOS;
KW antibody; purification; detection; inflammatory; psoriasis; uveitis;
KW rheumatoid arthritis.
OS Synthetic.
XX
PN WO9639858-A1.
XX
PD 19-DEC-1996.
XX
PF 03-JUN-1996; 96WO-US08447.
XX
PR 07-JUN-1995; 95US-0485075.
XX
PA (MERI) MERCK & CO INC.
XX
PI Calaycay JR, Mumford RA, Schmidt JA, Weidner JR;
XX
DR WPI; 1997-106622/10.

XX
PT Purifcn. of human inducible nitric oxide synthase from cultured
PT cells - using specific antibody directed against a C-terminal
PT epitope
XX
PS Example 3; Fig 2; 52pp; English.
XX
CC This is the epitope of the human inducible nitric oxide synthase (hiNOS)
CC without the C-terminal leucine. The primary wild type antigen is given
CC in AAW12317. It was shown that the C-terminal leucine of the epitope is
CC required for antibody recognition by the ability of radiolabelled
CC peptide to be displaced by unlabeled peptide, and by the specificity of
CC the antibody to a C-terminal leucine. Antibodies raised against the
CC primary antigen can be used to purify hiNOS from cultured cells in a
CC claimed method. Elevated levels of hiNOS are implicated in a wide variety
CC of diseases, particularly inflammatory conditions such as psoriasis,
CC uveitis, rheumatoid arthritis and many others. The antibodies can be
CC used to detect or quantify hiNOS (by usual immunoassay or
CC immunolocalisation techniques), also to evaluate the activity of iNOS
CC modulators (which are potentially useful as therapeutic agents).
XX
SQ Sequence 9 AA;

Qy 1 YRXXL 5
        | | |
Db 1 yrasl 5

Query Match   88.9%; Score 16; DB 18; Length 9;
Best Local Similarity   60.0%; Pred. No. 4.3e-05;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 2
ID AAW23311 standard; peptide; 10 AA.
XX
AC AAW23311;
XX
DT 30-MAR-1998 (first entry)
XX
DE Human inducible nitric oxide synthase immunogenic peptide.
XX
KW Inducible nitric oxide synthase; iNOS; endothelial nitric oxide synthase;
KW eNOS; diagnosis; detection; inflammatory response.
XX
OS Synthetic.
OS Homo sapiens.
XX
PN WO9734014-A1.
XX
PD 18-SEP-1997.
XX
PF 14-MAR-1997; 97WO-US04293.
XX
PR 15-MAR-1996; 96US-0013470.
XX
PA (UIYA) UNIV YALE.
XX
PI Sessa WC, Smith SD, Weiss RM, Wheeler MA;
XX
DR WPI; 1997-470090/43.
XX
PT Detection of inducible nitric oxide synthase - useful for monitoring
PT levels over periods of time, or to diagnose disease associated with
PT iNOS levels
XX
PS Disclosure; Page 9; 32pp; English.
XX
CC This sequence is an immunogenic peptide for detecting inducible nitric
CC oxide synthase (iNOS) immunoreactivity which is used in a novel method
CC to detect iNOS in a biological sample. The level of iNOS is related to
CC an inflammatory response, e.g. infection mediated by bacteria, yeast or

CC a virus, organ transplant rejection, rheumatoid arthritis, interstitial
CC cystitis or cancer. By monitoring levels of iNOS over a period of time
CC it is possible to diagnose a disease or condition associated with
CC increased or decreased iNOS.
XX
SQ Sequence 10 AA;

Qy 1 YRXXL 5
        | | |
Db 1 yrasl 5

Query Match   88.9%; Score 16; DB 18; Length 10;
Best Local Similarity   60.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 3
ID AAW12319 standard; peptide; 10 AA.
XX
AC AAW12319;
XX
DT 20-OCT-1997 (first entry)
XX
DE Radiolabelled HiNOS C-terminal immunogenic peptide.
XX
KW C-terminal; epitope; human; inducible nitric oxide synthase; hiNOS;
KW antibody; purification; detection; inflammatory; psoriasis; uveitis;
KW rheumatoid arthritis.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FH
FT Modified-site   1
FT   /note= "125-I labeled Tyrosine"
FT Peptide   4..10
FT   /note= "primary antigen"
FT Modified-site   10
FT   /note= "required for antibody recognition; optionally
FT   amidated in sensitivity assay"
XX
PN WO9639858-A1.
XX
PD 19-DEC-1996.
XX
PF 03-JUN-1996; 96WO-US08447.
XX
PR 07-JUN-1995; 95US-0485075.
XX
PA (MERI) MERCK & CO INC.
XX
PI Calaycay JR, Mumford RA, Schmidt JA, Weidner JR;
XX
DR WPI; 1997-108622/10.
XX
PT Purifcn. of human inducible nitric oxide synthase from cultured
PT cells - using specific antibody directed against a C-terminal
PT epitope
XX
PS Example 3; Page 19; 52pp; English.
XX
CC This sequence is a synthetic radio-iodinated peptide based on a C-
CC terminal epitope of human inducible nitric oxide synthase (hiNOS). It
CC was used to generate a standard curve in a competitive radioimmunoassay
CC for hiNOS products. It was shown that the C-terminal leucine is required
CC for antibody recognition by the ability of the radiolabelled peptide to
CC be displaced by unlabeled peptide. The specificity of the antibody to a
CC C-terminal leucine was also shown with use of peptides having amino acid
CC substitutions/modifications at the C-terminus (see AAW12320-22).
CC Antibodies raised against the primary antigen can be used to purify hiNOS
CC from cultured cells in a claimed method. Elevated levels of hiNOS are
CC implicated in a wide variety of diseases, particularly inflammatory

```
                    GenCore version 4.5
              Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           April 21, 2002, 10:15:54 ; Search time 94.43 Seconds
                                              (without alignments)
                                              1.192 Million cell updates/sec

Title:
Perfect score:    18
Sequence:         1 YRXXL 5

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters:    212252

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :       Issued_Patents_AA:*
                 1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                 2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                 3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                 4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                 5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	16	88.9	13	4	US-09-461-697-318	Sequence 318, App
2	16	88.9	30	4	US-09-461-697-314	Sequence 314, App
3	16	88.9	32	4	US-08-606-143-38/	Sequence 38, Appl
4	16	88.9	35	2	US-08-807-332B-9	Sequence 9, Appli
5	16	88.9	35	4	US-09-338-876-9	Sequence 9, Appli
6	16	88.9	40	1	US-08-735-963-3	Sequence 3, Appli
7	16	88.9	40	2	US-09-105-057-3	Sequence 3, Appli
8	16	88.9	40	2	US-09-304-214-3	Sequence 3, Appli
9	16	88.9	48	2	US-08-637-759B-480	Sequence 480, App
10	16	88.9	48	3	US-08-871-355A-480	Sequence 480, App
11	16	88.9	113	4	US-09-284-033-7	Sequence 7, Appli
12	16	88.9	113	4	US-08-729-834B-7	Sequence 7, Appli
13	16	88.9	119	4	US-09-046-479-4	Sequence 4, Appli
14	16	88.9	190	4	US-08-606-143-37	Sequence 37, Appl
15	16	88.9	192	1	US-08-735-963-2	Sequence 2, Appli
16	16	88.9	192	2	US-09-105-057-2	Sequence 2, Appli
17	16	88.9	192	2	US-09-304-214-2	Sequence 2, Appli
18	16	88.9	200	2	US-08-606-143-42	Sequence 42, Appl
19	16	88.9	208	2	US-08-606-143-40	Sequence 40, Appl
20	16	88.9	208	2	US-08-606-143-41	Sequence 41, Appl
21	16	88.9	208	2	US-08-606-143-43	Sequence 43, Appl
22	16	88.9	240	1	US-08-261-822A-82	Sequence 82, Appl
23	16	88.9	240	5	PCT-US95-07744A-82	Sequence 82, App
24	16	88.9	243	4	US-09-188-930-295	Sequence 295, App
25	16	88.9	243	4	US-09-140-804-2	Sequence 2, Appli
26	16	88.9	245	2	US-08-685-992-26	Sequence 26, Appl
27	16	88.9	245	2	US-09-144-925-26	Sequence 26, Appl
28	16	88.9	310	4	US-09-166-412-4	Sequence 4, Appli
29	16	88.9	316	1	US-08-038-932B-1	Sequence 1, Appli
30	16	88.9	316	1	US-08-656-349-1	Sequence 1, Appli
31	16	88.9	316	1	US-08-682-643-4	Sequence 4, Appli
32	16	88.9	316	3	US-09-104-623A-5	Sequence 5, Appli
33	16	88.9	316	4	US-09-415-751-35	Sequence 35, Appl
34	16	88.9	362	2	US-08-282-197C-60	Sequence 60, Appl
35	16	88.9	392	4	US-09-046-992-4	Sequence 4, Appli
36	16	88.9	396	2	US-08-391-259-2	Sequence 2, Appli
37	16	88.9	420	1	US-08-391-259-7	Sequence 7, Appli
38	16	88.9	420	1	US-08-391-259-10	Sequence 10, Appl
39	16	88.9	420	1	US-08-391-259-11	Sequence 11, Appl
40	16	88.9	420	2	US-08-839-425-2	Sequence 2, Appli
41	16	88.9	420	2	US-08-839-425-7	Sequence 7, Appli
42	16	88.9	420	2	US-08-839-425-10	Sequence 10, Appl
43	16	88.9	420	2	US-08-839-425-11	Sequence 11, Appl
44	16	88.9	456	1	US-08-021-601-6	Sequence 6, Appli
45	16	88.9	456	1	US-08-082-849B-6	Sequence 6, Appli
46	16	88.9	456	1	PCT-US94-01624-6	Sequence 6, Appli
47	16	88.9	462	2	US-08-484-438-42	Sequence 42, Appl
48	16	88.9	472	1	US-08-021-601-8	Sequence 8, Appli
49	16	88.9	472	1	US-08-082-849B-8	Sequence 8, Appli
50	16	88.9	472	5	PCT-US94-01624-8	Sequence 8, Appli

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

ALIGNMENTS

```
RESULT 1
US-09-461-697-318
; Sequence 318, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT APPLICATION NUMBER: US/09/461,697
; CURRENT FILING DATE: 1999-12-14
; NUMBER OF SEQ ID NOS: 466
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 318
; LENGTH: 13
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-318

Query Match             88.9%;  Score 16; DB 4; Length 13;
Best Local Similarity   60.0%;  Pred. No. 73;
Matches    3; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy          1 YRXXL 5
              | || |
Db          2 YRSSL 6

RESULT 2
US-09-461-697-314
; Sequence 314, Application US/09461697
; Patent No. 6277974
; GENERAL INFORMATION:
; APPLICANT: COGENT NEUROSCIENCE, Inc.
; APPLICANT: Lo, Donald C.
; APPLICANT: Barney, Shawn
; APPLICANT: Thomas, Mary Beth
```

```
; APPLICANT: Portbury, Stuart D.
; APPLICANT: Puranam, Kasturi
; APPLICANT: Katz, Lawrence C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR DIAGNOSING
; TITLE OF INVENTION: AND TREATING CONDITIONS, DISORDERS, OR DISEASES INVOLVING
; TITLE OF INVENTION: CELL DEATH
; FILE REFERENCE: 10001-005-999
; CURRENT FILING DATE: 1999-12-14
; CURRENT APPLICATION NUMBER: US/09/461,697
; SOFTWARE: FastSEQ for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 466
; SEQ ID NO 314
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-697-314

Query Match      88.9%;  Score 16;  DB 4;  Length 30;
Best Local Similarity  60.0%;  Pred. No. 1.7e+02;
Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 YRXXL 5
        | | |
Db   19 YRSSL 23

RESULT  3
US-08-606-143-38
Sequence 38, Application US/0860143
Patent No. 5856155
; GENERAL INFORMATION:
; APPLICANT: Li, Min
; TITLE OF INVENTION: COMPOUNDS AND RELATED METHODS FOR
; TITLE OF INVENTION: MODULATING POTASSIUM ION CHANNELS AND ASSAYS FOR SUCH
; TITLE OF INVENTION: COMPOUNDS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Leydig, Voit & Mayer, Ltd.
; STREET: Two Prudential Plaza, Suite 4900
; CITY: Chicago
; STATE: IL
; COUNTRY: USA
; ZIP: 60601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/606,143
; FILING DATE: 23-FEB-1996
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kilyk Jr., John
; REGISTRATION NUMBER: 30763
; REFERENCE/DOCKET NUMBER: 71756
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (312) 616-5600
; TELEFAX: (312) 616-5700
; TELEX: 25-3533
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 32 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-606-143-38

Query Match      88.9%;  Score 16;  DB 2;  Length 32;
Best Local Similarity  60.0%;  Pred. No. 1.8e+02;
```

```
Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 YRXXL 5
        | | |
Db   16 YRSTL 20

RESULT  4
US-08-807-332B-9
Sequence 9, Application US/08807332B
Patent No. 5959074
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; TITLE OF INVENTION: GENE RECOMBINATION
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
; STATE: CO
; COUNTRY: U.S.A.
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,332B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kovarik, Joseph E.
; REGISTRATION NUMBER: 33,005
; REFERENCE/DOCKET NUMBER: 2879-39
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/863-9700
; TELEFAX: 303/863-0223
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 35 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-807-332B-9

Query Match      88.9%;  Score 16;  DB 2;  Length 35;
Best Local Similarity  60.0%;  Pred. No. 1.9e+02;
Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy    1 YRXXL 5
        | | |
Db   30 YRATL 34

RESULT  5
US-09-338-876-9
Sequence 9, Application US/09338876
Patent No. 6187584
; GENERAL INFORMATION:
; APPLICANT: Dreyfus, David H.
; APPLICANT: Gelfand, Erwin W.
; TITLE OF INVENTION: PRODUCTS AND PROCESSES FOR REGULATION OF
; TITLE OF INVENTION: GENE RECOMBINATION
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross
; STREET: 1700 Lincoln St., Suite 3500
; CITY: Denver
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:17:56 ; Search time 102.84 Seconds
                                   (without alignments)
                                   5.926 Million cell updates/sec

Title: US-09-281-760C-3
Perfect score: 24
Sequence: 1 LXXYRXXL 8

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   PIR_68:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	Match ID	Description
1	20	83.3	384	2	S34816	nitrogenase cofact
2	20	83.3	412	2	T21419	hypothetical prote
3	19	79.2	152	2	T44544	conserved hypothet
4	19	79.2	208	2	T36390	hypothetical prote
5	19	79.2	226	2	G81894	probable integral
6	19	79.2	226	2	D81125	hypothetical prote
7	19	79.2	258	2	H83373	hypothetical prote
8	19	79.2	324	2	T09873	probable cellulase
9	19	79.2	328	2	B75403	minicell-associate
10	19	79.2	354	2	B65219	phnI protein - Esc
11	19	79.2	354	2	G86104	phosphonate metabo
12	19	79.2	366	2	D83224	conserved hypothet
13	19	79.2	377	2	T05354	hypothetical prote
14	19	79.2	377	2	C83372	hypothetical prote
15	19	79.2	401	2	S16693	hypothetical prote
16	19	79.2	440	2	T47906	FUSCA PROTEIN FUS6
17	19	79.2	487	2	T32941	hypothetical prote
18	19	79.2	508	2	T50180	nucleolar protein
19	19	79.2	545	1	RRBPBM	RNA-directed RNA p
20	19	79.2	566	2	S35231	hoxX protein - Bra
21	19	79.2	581	2	B72221	hypothetical prote
22	19	79.2	597	2	T10034	hypothetical prote
23	19	79.2	601	2	T14752	hypothetical prote
24	19	79.2	710	2	S26606	hypothetical prote
25	19	79.2	790	2	D83011	conserved hypothet
26	19	79.2	909	2	A73337	probable importin
27	19	79.2	1067	2	T39449	collagen alpha 1(x
28	19	79.2	1388	2	A53317	protein T23E23.4 [
29	18	75.0	48	2	B86373	nifG protein - Esc
30	18	75.0	69	2	S44888	hypothetical prote
31	18	75.0	77	2	C70967	hypothetical prote
32	18	75.0	96	2	T17209	hypothetical prote
33	18	75.0	98	2	S35469	NADH dehydrogenase
34	18	75.0	102	2	C95595	hypothetical prote
35	18	75.0	103	2	T24311	hypothetical prote
36	18	75.0	106	2	T48070	hypothetical prote
37	18	75.0	109	2	C85718	probable tail comp
38	18	75.0	109	2	C85742	probable tail comp
39	18	75.0	109	2	C85817	hypothetical prote
40	18	75.0	113	2	T14406	hypothetical prote
41	18	75.0	119	1	MSPG	motilin precursor
42	18	75.0	133	1	S24315	motilin precursor
43	18	75.0	144	2	G70023	hypothetical prote
44	18	75.0	166	2	G83176	hypothetical prote
45	18	75.0	175	2	T10067	hypothetical prote
46	18	75.0	176	2	F71321	conserved hypothet
47	18	75.0	190	2	B42957	iron-sulfur protei
48	18	75.0	192	2	F70359	hypothetical prote
49	18	75.0	197	2	T47159	hypothetical prote
50	18	75.0	198	1	A57987	2K112.6 protein -

ALIGNMENTS

RESULT 1
S34816
nitrogenase cofactor synthesis protein nifS - Rhodobacter capsulatus
N;Contains: L-cysteine sulfurtransferase
C;Species: Rhodobacter capsulatus
C;Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 07-Dec-1999
C;Accession: S34816
R;Masepohl, B.; Angermueller, S.; Hennecke, S.; Huebner, P.; Moreno-Vivian, C.; Klipp
Mol. Gen. Genet. 238, 369-382, 1993
A;Title: Nucleotide sequence and genetic analysis of the Rhodobacter capsulatus ORF6-
A;Reference number: S34814; MUID:93261420
A;Accession: S34816
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-384 <MAS>
A;Cross-references: EMBL:X68444; NID:g297915; PIDN:CAA48487.1; PID:g297918
C;Genetics:
A;Gene: nifS
C;Superfamily: nitrogen fixation protein nifS
C;Keywords: phosphoprotein; pyridoxal phosphate; sulfurtransferase
F;203/binding site: pyridoxal phosphate (Lys) (covalent) #status predicted
F;325/active site: Cys (cysteine persulfide intermediate) #status predicted

Query Match   83.3%; Score 20; DB 2; Length 384;
Best Local Similarity 50.0%; Pred. No. 3e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

Qy   1 LXXYRXXL 8
          | || |
Db 133 LAAYRAL 140

RESULT 2
T21419
hypothetical protein F26E4.13 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C;Accession: T21419
R;Lightning, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19419
A;Accession: T21419
A;Molecule type: DNA
A;Residues: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-412 <WIL>
A;Cross-references: EMBL:Z81070; PIDN:CAB03005.1; GSPDB:GN00019; CESP:F26E4.13

A;Experimental source: clone F26E4
C;Genetics:
A;Gene: CESP:F26E4.13
A;Map position: 1
A;Introns: 359/2; 400/3

```
Qy     1 LXXYRXXL 8
           | ||  |
Db    50 LATYRTAL 57
```

Query Match   83.3%;   Score 20; DB 2; Length 412;
Best Local Similarity   50.0%; Pred. No. 3.3e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

RESULT 3
T44544
conserved hypothetical protein PA0621 [imported] - Pseudomonas aeruginosa
C;Species: Pseudomonas aeruginosa
C;Date: 31-Jan-2000 #sequence_revision 31-Jan-2000 #text_change 31-Dec-2000
C;Accession: T44544; C83567
R;Nakayama, K.; Takashima, K.; Ishihara, H.; Shinomiya, T.; Kageyama, M.; Kanaya, S.; Oh
submitted to the EMBL Data Library, August 1999
A;Description: Genetic relationship between bacteriocins and bacteriophages.
A;Reference number: Z22790
A;Accession: T44544
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-152 <NAK>
A;Cross-references: EMBL:AB030825; PIDN:BAA83159.1
A;Experimental source: strain PAO1
R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrener, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A;Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic patho
A;Reference number: A82950; MUID:20437337
A;Accession: C83567
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-152 <STO>
A;Cross-references: GB:AE004498; GB:AE004091; NID:g9946491; PIDN:AAG04010.1; GSPDB:GN001
A;Experimental source: strain PAO1
A;Gene: PA0621

```
Qy     1 LXXYRXXL 8
           | ||  |
Db   117 LMSYRASL 124
```

Query Match   79.2%;   Score 19; DB 2; Length 152;
Best Local Similarity   50.0%; Pred. No. 2.4e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

RESULT 4
T36390
hypothetical protein SCE94.30c - Streptomyces coelicolor
C;Species: Streptomyces coelicolor
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 03-Dec-1999
C;Accession: T36390
R;Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, April 1999
A;Reference number: Z21573
A;Accession: T36390
A;Molecule type: DNA
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Residues: 1-208 <OLI>
A;Cross-references: EMBL:AL049628; PIDN:CAB40879.1; GSPDB:GN0070; SCOEDB:SCE94.30c
A;Experimental source: strain A3(2)
C;Genetics:
A;Gene: SCOEDB:SCE94.30c

Query Match   79.2%;   Score 19; DB 2; Length 208;
Best Local Similarity   50.0%; Pred. No. 3.3e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1 LXXYRXXL 8
           | ||  |
Db    67 LARYRAAL 74
```

RESULT 5
G81894
probable integral membrane protein NMA1265 [imported] - Neisseria meningitidis
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 02-Feb-2001
C;Accession: G81894
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Mo
; Holroyd, S.; Jagels, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandre
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491
A;Reference number: A81775; MUID:20222556
A;Accession: G81894
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-226 <PAR>
A;Cross-references: GB:AL162755; GB:AL157959; NID:g7379742; PIDN:CAB84520.1; PID:g737
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA1265

Query Match   79.2%;   Score 19; DB 2; Length 226;
Best Local Similarity   50.0%; Pred. No. 3.5e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1 LXXYRXXL 8
           | ||  |
Db    35 LARYRSAL 42
```

RESULT 6
D81125
hypothetical protein NMB1066 [imported] - Neisseria meningitidis (strain MC58 serogro
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 19-Jan-2001
C;Accession: D81125
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen,
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.;
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: A81000; MUID:20175755
A;Accession: D81125
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-226 <TET>
A;Cross-references: GB:AE002457; GB:AE002098; NID:g7226304; PIDN:AAF41462.1; PID:g722
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1066

Query Match   79.2%;   Score 19; DB 2; Length 226;
Best Local Similarity   50.0%; Pred. No. 3.5e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

```
Qy     1 LXXYRXXL 8
           | ||  |
Db    35 LARYRSAL 42
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:37:23 ; Search time 85.5 Seconds
                 (without alignments)
                 3.431 Million cell updates/sec

Title: US-09-281-760C-3
Perfect score: 24
Sequence: 1 LXXYRXXL 8
Scoring table:
Run on:

Database :   SwissProt_39:*

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	20	83.3	384	1	NIFS_RHOCA	Q07177 rhodobacter
2	20	83.3	412	1	RNC_CAEEL	O01326 caenorhabdi
3	19	79.2	38	1	CPRP_CANPG	P81033 cancer pagu
4	19	79.2	208	1	RL13_CHLSW	O48513 chlamydomon
5	19	79.2	263	1	YEMT_HUMAN	O95050 homo sapien
6	19	79.2	354	1	PHNI_ECOLI	P16687 escherichia
7	19	79.2	368	1	PHNI_RHIME	Q52986 rhizobium m
8	19	79.2	441	1	FUS6_ARATH	P45432 arabidopsis
9	19	79.2	544	1	RRPO_BPMS2	P05585 bacteriopha
10	19	79.2	566	1	HOXX_BRAJA	P31907 bradyrhizob
11	19	79.2	1004	1	EPA8_MOUSE	O09127 mus musculu
12	19	79.2	1067	1	IMB4_SCHPO	O60100 schizosacch
13	19	79.2	1388	1	CA1E_HUMAN	P39059 homo sapien
14	18	75.0	69	1	YOG6_CAEEL	P34615 caenorhabdi
15	18	75.0	92	1	MOTI_HORSE	O46617 equus cabal
16	18	75.0	98	1	NULM_CROLA	P34193 crossostoma
17	18	75.0	113	1	YX21_CAEEL	Q11192 caenorhabdi
18	18	75.0	115	1	MOTI_MACMU	O18811 macaca mula
19	18	75.0	116	1	MOTI_FELCA	Q9xse2 felis silve
20	18	75.0	119	1	MOTI_PIG	P01307 sus scrofa
21	18	75.0	130	1	UCR6_SCHMA	O01374 schistosoma
22	18	75.0	133	1	MOTI_RABIT	P27114 oryctolagus
23	18	75.0	190	1	COOF_RHORU	P31894 rhodospiril
24	18	75.0	192	1	Y678_AQUAE	O66906 aquifex aeo
25	18	75.0	195	1	NRTN_MOUSE	P97463 mus musculu
26	18	75.0	197	1	NRTN_HUMAN	Q99748 homo sapien
27	18	75.0	198	1	NRFG_ECOLI	P32712 escherichia
28	18	75.0	214	1	ET3_MOUSE	P48299 mus musculu
29	18	75.0	264	1	NNMT_MOUSE	O55239 mus musculu
30	18	75.0	298	1	XERD_SALTY	P55889 salmonella
31	18	75.0	298	1	YSMK_CAEEL	Q19408 caenorhabdi
32	18	75.0	325	1	DBP_HUMAN	Q10586 homo sapien
33	18	75.0	325	1	DBP_MOUSE	Q60925 mus musculu
34	18	75.0	325	1	DBP_RAT	P16443 rattus norv
35	18	75.0	340	1	ENDR_PAEPO	P27871 paenibacill
36	18	75.0	355	1	FKB8_HUMAN	Q14318 homo sapien
37	18	75.0	355	1	FKB8_MOUSE	O35465 mus musculu
38	18	75.0	364	1	DRG_HUMAN	P55039 homo sapien
39	18	75.0	364	1	DRG2_MOUSE	Q9qxb9 mus musculu
40	18	75.0	393	1	YWCF_BACSU	P39604 bacillus su
41	18	75.0	419	1	CTEL_MOUSE	O55137 mus musculu
42	18	75.0	419	1	CTEL_RAT	O82067 rattus norv
43	18	75.0	421	1	PTE2_HUMAN	Q9qyr7 homo sapien
44	18	75.0	432	1	PTE2_MOUSE	P49753 mus musculu
45	18	75.0	432	1	Y412_ARATH	O04658 arabidopsis
46	18	75.0	439	1	SUN_HAEIN	P44788 haemophilus
47	18	75.0	453	1	MTEI_MOUSE	Q9qyr9 mus musculu
48	18	75.0	453	1	MTEI_RAT	O55171 rattus norv
49	18	75.0	461	1	MYH_SCHPO	Q10159 schizosacch
50	18	75.0	478	1	HLVD_PASHA	P16534 pasteurella

ALIGNMENTS

RESULT 1
NIFS_RHOCA

```
ID   NIFS_RHOCA   STANDARD;   PRT;   384 AA.
AC   Q07177;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   CYSTEINE DESULFURASE (EC 4.4.1.-) (NITROGENASE METALLOCLUSTERS
DE   BIOSYNTHESIS PROTEIN NIFS).
GN   NIFS.
OS   Rhodobacter capsulatus (Rhodopseudomonas capsulata).
OC   Bacteria; Proteobacteria; alpha subdivision; Rhodobacter group;
OC   Rhodobacter.
OX   NCBI_TaxID=1061;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B10;
RX   MEDLINE=93261420; PubMed=8492805;
RA   Masepohl B., Angermueller S., Hennecke S., Huebner P.,
RA   Moreno-Vivian C., Klipp W.;
RT   "Nucleotide sequence and genetic analysis of the Rhodobacter
RT   capsulatus ORF6-nifUI SVW gene region: possible role of NifW in
RT   homonitrate processing.";
RL   Mol. Gen. Genet. 238:369-382(1993).
CC   -!- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR ATOMS FROM
CC       CYSTEINE TO PRODUCE ALANINE. SEEMS TO PARTICIPATE IN THE
CC       BIOSYNTHESIS OF THE NITROGENASE METALLOCLUSTERS BY PROVIDING THE
CC       INORGANIC SULFUR REQUIRED FOR THE FE-S CORE FORMATION (BY
CC       SIMILARITY).
CC   -!- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC   -!- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC   -!- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT
CC       AMINOTRANSFERASES. NIFS/ISCS SUBFAMILY.
CC   -------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -------------------------------------------------------------------------
DR   EMBL; X68444; CAA48487.1; -.
DR   PIR; S34816; S34816.
DR   Interpro; IPR000192; Aminotransf_class_V.
DR   Pfam; PF00266; aminotran_5; 1.
DR   PROSITE; PS00595; AA_TRANSFER_CLASS_5; 1.
KW   Nitrogen fixation; Lyase; Pyridoxal phosphate.
FT   BINDING     203    203    PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT   ACT_SITE    325    325    BY SIMILARITY.
SQ   SEQUENCE   384 AA;   40282 MW;   C2C4A282D80A8391 CRC64;
```

RESULT 2

```
Qy      1 LXXYRXXL 8
          | ||| |
Db    133 LAAYRAAL 140
```

Query Match   83.3%; Score 20; DB 1; Length 384;
Best Local Similarity   50.0%; Pred. No. 1.1e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

ID   RNC_CAEEL   STANDARD;   PRT;   412 AA.
AC   O01326;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PUTATIVE RIBONUCLEASE III (EC 3.1.26.3) (RNASE III).
GN   F26E4.13.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BRISTOL N2;
RA   Lightning J.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
CC   -!- CATALYTIC ACTIVITY: ENDONUCLEOLYTIC CLEAVAGE TO 5'-PHOSPHO-
CC   MONOESTER.
CC   -!- SIMILARITY: BELONGS TO THE RNASE III FAMILY.
CC   -!- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; Z81070; CAB03005.1; -.
DR   WormPep; F26E4.13; CE09697.
DR   InterPro; IPR000159; DS_RBD.
DR   InterPro; IPR000999; RNase_3.
DR   Pfam; PF00035; dsrm; 1.
DR   Pfam; PF00636; Ribonuclease_3; 2.
DR   SMART; SM00358; DSRM; 1.
DR   SMART; SM00535; RIBOc; 2.
DR   PROSITE; PS50137; DS_RBD; 1.
DR   PROSITE; PS00517; RNASE_3_1; 2.
DR   PROSITE; PS50142; RNASE_3_2; 2.
KW   Hypothetical protein; Hydrolase; Nuclease; Endonuclease.
FT   DOMAIN   313   388   DRBM.
SQ   SEQUENCE 412 AA; 47087 MW; E084081AB87AF84 CRC64;
```

RESULT 3

```
Qy      1 LXXYRXXL 8
          | ||| |
Db     50 LATYRTAL 57
```

Query Match   83.3%; Score 20; DB 1; Length 412;
Best Local Similarity   50.0%; Pred. No. 1.1e+02;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

ID   CPRP_CANPG   STANDARD;   PRT;   38 AA.
AC   P81033;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   15-DEC-1998 (Rel. 37, Last annotation update)
DE   CHH PRECURSOR RELATED PEPTIDE (CPRP).
OS   Cancer pagurus (Rock crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Cancroidea; Cancridae; Cancer.
OX   NCBI_TaxID=6755;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Sinus gland;
RX   MEDLINE=99025664; PubMed=9809792;
RA   Chung J.S., Wilkinson M.C., Webster S.G.;
RT   "Amino acid sequences of both isoforms of crustacean hyperglycemic
RT   hormone (CHH) and corresponding precursor-related peptide in Cancer
RT   pagurus.";
RL   Regul. Pept. 77:17-24(1998).
CC   -!- TISSUE SPECIFICITY: PRODUCED BY THE MEDULLA TERMINALIS X-ORGAN IN
CC   THE EYESTALKS AND TRANSPORTED TO THE SINUS GLAND WHERE IT IS
CC   STORED AND RELEASED.
KW   Neuropeptide; Hormone.
SQ   SEQUENCE 38 AA; 3969 MW; C979C87EE31ABB90 CRC64;
```

RESULT 4

```
Qy      1 LXXYRXXL 8
          | ||||| |
Db     13 LASYRGAL 20
```

Query Match   79.2%; Score 19; DB 1; Length 38;
Best Local Similarity   50.0%; Pred. No. 18;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;

ID   RL13_CHLSW   STANDARD;   PRT;   208 AA.
AC   O48513;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   60S RIBOSOMAL PROTEIN L13 (BBC1 PROTEIN HOMOLOG).
GN   RPL13 OR BBC1.
OS   Chlamydomonas sp. (strain W80).
OC   Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Volvocales;
OC   Chlamydomonadaceae; Chlamydomonas.
OX   NCBI_TaxID=103365;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Kanaboshi H., Ikeda K., Miyasaka H.;
RT   "Isolation of several anti-stress genes from the halotolerant green
RT   alga Chlamydomonas by a simple functional expression screening in E.
RT   coli.";
RL   Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
CC   -!- SIMILARITY: BELONGS TO THE L13E FAMILY OF RIBOSOMAL PROTEINS.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AB009086; BAA23724.1; -.
DR   MendeL; 27088; ChlSp:2745;27085.
DR   InterPro; IPR001380; Ribosomal_L13E.
DR   Pfam; PF01294; Ribosomal_L13e; 1.
DR   ProDom; PD004443; Ribosomal_L13E; 1.
DR   PROSITE; PS01104; RIBOSOMAL_L13E; 1.
KW   Ribosomal protein.
SQ   SEQUENCE 208 AA; 23545 MW; F9BB54E2550EA43B CRC64;
```

Query Match       79.2%;  Score 19;  DB 1;  Length 208;

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:           April 21, 2002, 10:35:46 ; Search time 191.29 Seconds
                  (without alignments)
                  6.117 Million cell updates/sec

Title:            US-09-281-760C-3
Perfect score:    24
Sequence:         1 LXXYRXXL 8

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         473505 seqs, 14627329 residues

Total number of hits satisfying chosen parameters:

Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Database :  SPTREMBL_17:*
   1:  sp_archea:*
   2:  sp_bacteria:*
   3:  sp_plant:*
   4:  sp_fungi:*
   5:  sp_human:*
   6:  sp_invertebrate:*
   7:  sp_mammal:*
   8:  sp_mhc:*
   9:  sp_organelle:*
  10:  sp_phage:*
  11:  sp_rodent:*
  12:  sp_virus:*
  13:  sp_vertebrate:*
  14:  sp_unclassified:*                          473505

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Query Match | Length | DB | ID | Description |
|---|---|---|---|---|---|---|
| 1 | 20 | 83.3 | 315 | 4 | Q9H4J8 | homo sapien |
| 2 | 20 | 83.3 | 331 | 4 | O15405 | homo sapien |
| 3 | 20 | 83.3 | 490 | 2 | Q9F2U5 | streptomyce |
| 4 | 20 | 83.3 | 526 | 4 | O94900 | homo sapien |
| 5 | 20 | 83.3 | 656 | 12 | Q9DWC3 | rat cytomeg |
| 6 | 20 | 83.3 | 860 | 5 | Q9U9Q8 | caenorhabdi |
| 7 | 19 | 79.2 | 139 | 2 | Q9RDQ0 | streptomyce |
| 8 | 19 | 79.2 | 152 | 2 | Q9S575 | streptomyce |
| 9 | 19 | 79.2 | 197 | 2 | S53172 | pseudomonas |
| 10 | 19 | 79.2 | 208 | 2 | Q9X8N5 | rhodobacter |
| 11 | 19 | 79.2 | 226 | 2 | Q9JZG5 | streptomyce |
| 12 | 19 | 79.2 | 226 | 2 | Q9X8N5 | neisseria m |
| 13 | 19 | 79.2 | 258 | 2 | Q9JUL0 | neisseria m |
| 14 | 19 | 79.2 | 258 | 10 | Q9IU7 | pseudomonas |
| 15 | 19 | 79.2 | 263 | 4 | Q9XIC2 | arabidopsis |
| 16 | 19 | 79.2 | 263 | 4 | Q9UHQ0 | homo sapien |
| 17 | 19 | 79.2 | 263 | 6 | Q9UBY4 | homo sapien |
| 18 | 19 | 79.2 | 324 | 10 | O23954 | gossypium h |
| 19 | 19 | 79.2 | 328 | 2 | Q9RUL5 | deinococcus |
| 20 | 19 | 79.2 | 366 | 2 | Q9HYM3 | pseudomonas |
| 21 | 19 | 79.2 | 377 | 2 | Q9IIT4 | pseudomonas |
| 22 | 19 | 79.2 | 377 | 10 | Q9SUU8 | arabidopsis |
| 23 | 19 | 79.2 | 401 | 2 | O48910 | mycobacteri |
| 24 | 19 | 79.2 | 403 | 2 | O52629 | paracoccus |
| 25 | 19 | 79.2 | 438 | 10 | Q9FGB9 | arabidopsis |
| 26 | 19 | 79.2 | 440 | 10 | Q9M2E9 | arabidopsis |
| 27 | 19 | 79.2 | 441 | 10 | Q9C5F0 | arabidopsis |
| 28 | 19 | 79.2 | 487 | 5 | O45012 | caenorhabdi |
| 29 | 19 | 79.2 | 508 | 3 | Q9P7S7 | schizosacch |
| 30 | 19 | 79.2 | 523 | 8 | O18432 | geodia cydo |
| 31 | 19 | 79.2 | 525 | 12 | O89175 | rice grassy |
| 32 | 19 | 79.2 | 525 | 12 | Q9QNA1 | rice grassy |
| 33 | 19 | 79.2 | 525 | 9 | Q9ELT1 | rice grassy |
| 34 | 19 | 79.2 | 526 | 9 | Q9T1C5 | enterobacte |
| 35 | 19 | 79.2 | 566 | 5 | Q9W243 | drosophila |
| 36 | 19 | 79.2 | 581 | 2 | Q9X230 | thermotoga |
| 37 | 19 | 79.2 | 597 | 2 | O33086 | mycobacteri |
| 38 | 19 | 79.2 | 597 | 2 | Q9CDD7 | mycobacteri |
| 39 | 19 | 79.2 | 659 | 10 | Q9FV60 | zea mays (m |
| 40 | 19 | 79.2 | 693 | 8 | Q9TK79 | calyptidiu |
| 41 | 19 | 79.2 | 706 | 2 | Q9A2Q2 | caulobacter |
| 42 | 19 | 79.2 | 710 | 8 | Q35064 | marchantia |
| 43 | 19 | 79.2 | 720 | 12 | Q9DT76 | tt virus. o |
| 44 | 19 | 79.2 | 732 | 8 | Q9TJS3 | pentas lanc |
| 45 | 19 | 79.2 | 748 | 12 | Q9DT81 | tt virus. o |
| 46 | 19 | 79.2 | 790 | 2 | Q9HU93 | pseudomonas |
| 47 | 19 | 79.2 | 853 | 4 | Q9P269 | homo sapien |
| 48 | 19 | 79.2 | 859 | 2 | Q33828 | thiobacillu |
| 49 | 19 | 79.2 | 909 | 2 | Q9RT44 | deinococcus |
| 50 | 19 | 79.2 | 963 | 2 | O06458 | thermus aqu |

ALIGNMENTS

RESULT 1
Q9H4J8

ID   Q9H4J8       PRELIMINARY;      PRT;   315 AA.
AC   Q9H4J8;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   DJ1108D11.2 (NOVEL HMG (HIGH MOBILITY GROUP) BOX PROTEIN SIMILAR TO
DE   KIAA0737, KIAA0808 AND TNRC9 (CAGF9)) (FRAGMENT).
GN   DJ1108D11.2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clark G.;
RL   Submitted (NOV-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AL034419; CAB81660.1; -.
DR   InterPro; IPR000135; Highmobly_12.
DR   InterPro; IPR000910; HMG_12_box.
DR   InterPro; IPR002965; P_rich_extensn.
DR   Pfam; PF00505; HMG_box; 1.
DR   PRINTS; PR00886; HIGHMOBLTY12.
DR   PRINTS; PR01217; PRICHEXTENSN.
DR   SMART; SM00398; HMG; 1.
FT   NON_TER       1.
FT   NON_TER       315.
SQ   SEQUENCE   315 AA;  33204 MW;  F77A3178A48A27D5 CRC64;

Query Match             83.3%;  Score 20;  DB 4;  Length 315;
Best Local Similarity   50.0%;  Pred. No. 7.7e+02;
Matches    4;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 LXXYRXXL 8
              | ||| |

Db  173 LAAYRASL 180

RESULT 2

ID  O15405    PRELIMINARY;    PRT;    331 AA.
AC  O15405;
DT  01-JAN-1998 (TrEMBLrel. 05, Created)
DT  01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT  01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE  CAGF9 (FRAGMENT).
GN  CAGF9.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=BRAIN;
RX  MEDLINE=97369492; PubMed=9225980;
RA  Margolis R.L., Abraham M.R., Gatchell S.B., Li S.H., Kidwai A.S.,
RA  Breschel T.S., Stine O.C., Callahan C., McInnis M.G., Ross C.A.;
RT  "cDNAs with long CAG trinucleotide repeats from human brain.";
RL  Hum. Genet. 100:114-122(1997).
DR  EMBL; U80736; AAB91435.1; -.
DR  HSSP; P07155; 1HMF.
DR  Interpro; IPR000135; Highmobility_12.
DR  Interpro; IPR000910; HMG_12_box.
DR  Pfam; PF00505; HMG_box; 1.
DR  PRINTS; PR00886; HIGHMOBLTY12.
DR  SMART; SM00398; HMG; 1.
FT  NON_TER  1
SQ  SEQUENCE  331 AA;  37161 MW;  71C00E4BFA10766B CRC64;

Qy  1  LXXYRXXL  8
       | | |  |
Db  74  LAAYRASL  81

Query Match     83.3%;    Score 20;   DB 4;   Length 331;
Best Local Similarity   50.0%;   Pred. No. 8.2e+02;
Matches  4;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

RESULT 3

ID  Q9F2U5    PRELIMINARY;    PRT;    490 AA.
AC  Q9F2U5;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE  PUTATIVE ATP-BINDING PROTEIN.
GN  SCD39.17c.
OS  Streptomyces coelicolor.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX  NCBI_TaxID=1902;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RA  Brown S.P., Harris D.;
RL  Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RA  Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL  Submitted (SEP-2000) to the EMBL/GenBank/DDBJ databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  STRAIN=A3(2);
RX  MEDLINE=97000351; PubMed=8843436;
RA  Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
RA  Kinashi H., Hopwood D.A.;
RT  "A set of ordered cosmids and a detailed genetic and physical map for
RT  the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL  Mol. Microbiol. 21:77-96(1996).
DR  EMBL; AL392146; CAC08270.1; -.
KW  ATP-binding.
SQ  SEQUENCE  490 AA;  53788 MW;  46BC07DAC1788FAB CRC64;

Qy  1  LXXYRXXL  8
       | | |  |
Db  374  LAAYRAAL  381

Query Match     83.3%;    Score 20;   DB 2;   Length 490;
Best Local Similarity   50.0%;   Pred. No. 1.2e+03;
Matches  4;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

RESULT 4

ID  O94900    PRELIMINARY;    PRT;    526 AA.
AC  O94900;
DT  01-MAY-1999 (TrEMBLrel. 10, Created)
DT  01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT  01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE  KIAA0808 PROTEIN.
GN  KIAA0808.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=BRAIN;
RX  MEDLINE=99087487; PubMed=9872452;
RA  Nagase T., Ishikawa K., Suyama M., Kikuno R., Miyajima N., Tanaka A.,
RA  Kotani H., Nomura N., Ohara O.;
RT  "Prediction of the coding sequences of unidentified human genes. XI.
RT  The complete sequences of 100 new cDNA clones from brain which code
RT  for large proteins in vitro.";
RL  DNA Res. 5:277-286(1998).
DR  EMBL; AB018351; BAA34528.1; -.
DR  HSSP; P07155; 1HMF.
DR  Interpro; IPR000135; Highmobility_12.
DR  Interpro; IPR000910; HMG_12_box.
DR  Pfam; PF00505; HMG_box; 1.
DR  PRINTS; PR00886; HIGHMOBLTY12.
DR  SMART; SM00398; HMG; 1.
SQ  SEQUENCE  526 AA;  57513 MW;  5AE4C5B07F2EB8FF CRC64;

Qy  1  LXXYRXXL  8
       | | |  |
Db  325  LAAYRASL  332

Query Match     83.3%;    Score 20;   DB 4;   Length 526;
Best Local Similarity   50.0%;   Pred. No. 1.3e+03;
Matches  4;  Conservative  0;  Mismatches  4;  Indels  0;  Gaps  0;

RESULT 5

ID  Q9DWC3    PRELIMINARY;    PRT;    656 AA.
AC  Q9DWC3;
DT  01-MAR-2001 (TrEMBLrel. 16, Created)
DT  01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT  01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE  PR77.
GN  PR77.
OS  Rat cytomegalovirus (strain Maastricht).
OC  Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC  Betaherpesvirinae; Muromegalovirus.
OX  NCBI_TaxID=79700;
RN  [1]

```
                          GenCore version 4.5
                     Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:         April 21, 2002, 10:14:08 ; Search time 158.02 Seconds
                                            (without alignments)
                                            3.750 Million cell updates/sec

Title:          US-09-281-760C-3
Perfect score:  24
Sequence:       1 LXXYRXXL 8
Scoring table:  BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched:       522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters:          522463

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

imum Match 0%
ximum Match 100%
Minimum DB seq length: 0
ximum DB seq length: 200000000

Database :   A_Geneseq_1101:*
   1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
   2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
   3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
   4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
   5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
   6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
   7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
   8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
   9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
  10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
  11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
  12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
  13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
  14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
  15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
  16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
  17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
  18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
  19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
  20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
  21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
  22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	20	83.3	388	22	AAM41453	Human colon cancer
2	20	83.3	576	22	AAB73519	Human transferase
3	19	79.2	42	22	AAM39667	Arabidopsis thalia
4	19	79.2	42	22	AAM14043	Arabidopsis thalia
5	19	79.2	42	22	AAM26449	Peptide #486 encod
6	19	79.2	176	21	AAM01783	Peptide #465 encod
7	19	79.2	180	22	AAW90875	Human HMW endostat
8	19	79.2	181	20	AAB49382	Human endostatin S
9	19	79.2	181	20	AAY22227	Human restin prote
10	19	79.2	191	21	AAW90876	Restin protein seq
11	19	79.2	219	21	AAG29713	Arabidopsis thalia

No.	Score	%	Length	DB	ID	Description
12	19	79.2	222	22	AAG74897	Human colon cancer
13	19	79.2	303	22	AAB73519	Human transferase
14	19	79.2	327	21	AAG29711	Arabidopsis thalia
15	19	79.2	377	21	AAG29712	Arabidopsis thalia
16	19	79.2	526	20	AAY33686	S. erythraea pFL37
17	19	79.2	963	17	AAR84082	Thermostable enzym
18	19	79.2	992	22	AAU00691	Ephrin type-A rece
19	19	79.2	1012	22	AAE04362	Human kinase (PKIN
20	18	75.0	11	12	AAR13244	Nocifensin anti-In
21	18	75.0	11	15	AAR48303	Antiinflammatory P
22	18	75.0	11	15	AAR48304	Antiinflammatory P
23	18	75.0	11	15	AAR48309	Antiinflammatory P
24	18	75.0	30	22	AAG98758	Antiinflammatory P
25	18	75.0	66	21	AAG55104	Human cell death P
26	18	75.0	69	22	AAG73544	Arabidopsis thalia
27	18	75.0	70	20	AAY19484	Human colon cancer
28	18	75.0	76	20	AAY16642	Amino acid sequenc
29	18	75.0	76	20	AAY16641	WO9914235 Seq ID N
30	18	75.0	89	21	AAG01676	Human secreted pro
31	18	75.0	91	22	AAM39497	Human polypeptide
32	18	75.0	95	20	AAY16643	Human secreted pro
33	18	75.0	95	22	AAY16644	WO9914235 Seq ID N
34	18	75.0	96	21	AAG00511	WO9914235 Seq ID N
35	18	75.0	96	21	AAG01677	Human secreted pro
36	18	75.0	96	22	AAM25654	Murine erythroid d
37	18	75.0	102	21	AAB15367	Amino acid sequenc
38	18	75.0	106	20	AAY37778	Arabidopsis thalia
39	18	75.0	106	21	AAG55843	Human protein sequ
40	18	75.0	110	20	AAW88842	Human secreted pro
41	18	75.0	110	20	AAY12157	Human secreted pro
42	18	75.0	117	20	AAG33864	Polypeptide fragme
43	18	75.0	117	22	AAP91380	Human 5' EST secre
44	18	75.0	119	10	AAP62651	Human 5' EST secre
45	18	75.0	119	22	AAB20102	Arabidopsis thalia
46	18	75.0	121	17	AAB20102	Porcine prepromoti
47	18	75.0	121	17	AAR97830	Motilin receptor,
48	18	75.0	121	21	AAR93611	Pig motilin, Sus
49	18	75.0	124	21	AAB38583	Kaposi's sarcoma a
50	18	75.0	141	21	AAG55842	Kaposi's sarcoma a
	18	75.0		21	AAG04689	Human secreted pro

ALIGNMENTS

RESULT   1
AAM41453

```
ID   AAM41453  standard; protein; 388 AA.
XX
AC   AAM41453;
XX
DT   22-OCT-2001  (first entry)
XX
DE   Human polypeptide SEQ ID NO 6384.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
XX
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.
XX
PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0552127,
PR   09-JUL-2000; 2000US-0598042,
PR   19-JUL-2000; 2000US-0620312.
```

PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
PA
XX
WPI; 2001-442253/47.
DR   N-PSDB; AAI60609.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 2; SEQ ID NO 6384; 10078pp; English.
XX

CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   388 AA;

Query Match   Best Local Similarity   83.3%;
                        Score 20; DB 22; Length 388;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;
                        Pred. No. 9.7e+02;

Qy   1 LXXYRXXL   8
           | || |
Db   131 laayrasl 138

RESULT   2
ID   AAM39667 standard; Protein; 576 AA.
AC   AAM39667;
XX
DT   22-OCT-2001 (first entry)
XX
DE   Human polypeptide SEQ ID NO 2812.
XX
KW   Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW   peripheral nervous system; neuropathy; central nervous system; CNS;
KW   Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW   amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW   chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW   leukaemia.
OS   Homo sapiens.
XX
PN   WO200153312-A1.
XX
PD   26-JUL-2001.
XX
PF   26-DEC-2000; 2000WO-US34263.

PR   21-JAN-2000; 2000US-0488725.
PR   25-APR-2000; 2000US-0553317.
PR   09-JUL-2000; 2000US-0598042.
PR   19-JUL-2000; 2000US-0620312.
PR   03-AUG-2000; 2000US-0653450.
PR   14-SEP-2000; 2000US-0662191.
PR   19-OCT-2000; 2000US-0693036.
PR   29-NOV-2000; 2000US-0727344.
XX
PA   (HYSE-) HYSEQ INC.
XX
PI   Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI   Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI   Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
PA
XX
WPI; 2001-442253/47.
DR   N-PSDB; AAI58823.
XX
PT   Novel nucleic acids and polypeptides, useful for treating disorders
PT   such as central nervous system injuries -
XX
PS   Example 4; SEQ ID NO 2812; 10078pp; English.
XX

CC   The invention relates to human nucleic acids (AAI57798-AAI61369) and
CC   the encoded polypeptides (AAM38642-AAM42213) with nootropic,
CC   immunosuppressant and cytostatic activity. The polynucleotides are useful
CC   in gene therapy. A composition containing a polypeptide or polynucleotide
CC   of the invention may be used to treat diseases of the peripheral nervous
CC   system, such as peripheral nervous injuries, peripheral neuropathy and
CC   localised neuropathies and central nervous system diseases, such as
CC   Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC   lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC   utilisation of the activities such as: Immune system suppression,
CC   Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC   and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC   assays for receptor activity, arthritis and inflammation, leukaemias and
CC   C.N.S disorders.
CC   Note: The sequence data for this patent did not form part of the printed
CC   specification.
XX
SQ   Sequence   576 AA;

Query Match   Best Local Similarity   83.3%;
                        Score 20; DB 22; Length 576;
Matches   4; Conservative   0; Mismatches   4; Indels   0; Gaps   0;
                        Pred. No. 1.4e+03;

Qy   1 LXXYRXXL   8
           | || |
Db   319 laayrasl 326

RESULT   3
ID   AAM14043 standard; Protein; 42 AA.
AC   AAM14043;
XX
DT   12-OCT-2001 (first entry)
XX
DE   Peptide #477 encoded by probe for measuring cervical gene expression.
XX
KW   Probe; human; microarray; gene expression; cervical cancer;
KW   cervical cancer.
OS   Homo sapiens.
XX
PN   WO200157278-A2.
XX
PD   09-AUG-2001.
XX
PF   30-JAN-2001; 2001WO-US00670.

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:15:56 ; Search time 94.43 Seconds
                                      (without alignments)
                                      1.906 Million cell updates/sec

Title: US-09-281-760C-3
Perfect score: 24
Sequence: 1 LXXYRXXL 8

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252
Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   Issued_Patents_AA:*
     1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
     2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
     3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
     4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
     5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
     6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	%	Query Match Length	DB	ID	Description
1	19	79.2	526	2	US-08-852-401-3	Sequence 3, Appli
2	19	79.2	963	1	US-08-537-002A-3	Sequence 3, Appli
3	19	79.2	963	3	US-08-863-010-3	Sequence 3, Appli
4	19	79.2	963	3	US-09-024-429-3	Sequence 3, Appli
5	18	75.0	11	6	5177060-30	Patent No. 5177060
6	18	75.0	30	4	US-09-461-697-314	Sequence 314, App
7	18	75.0	76	1	US-08-519-777-19	Sequence 19, Appl
8	18	75.0	76	1	US-08-519-777-22	Sequence 22, Appl
9	18	75.0	76	1	US-08-742-035-19	Sequence 19, Appl
10	18	75.0	76	1	US-08-742-035-22	Sequence 22, Appl
11	18	75.0	76	2	US-08-777-019-19	Sequence 19, Appl
12	18	75.0	76	2	US-08-777-019-22	Sequence 22, Appl
13	18	75.0	76	2	US-08-777-143-19	Sequence 19, Appl
14	18	75.0	76	2	US-08-777-143-22	Sequence 22, Appl
15	18	75.0	76	3	US-08-775-414-19	Sequence 19, Appl
16	18	75.0	76	3	US-08-775-414-22	Sequence 22, Appl
17	18	75.0	76	4	US-08-931-858E-19	Sequence 19, Appl
18	18	75.0	76	4	US-08-931-858E-22	Sequence 22, Appl
19	18	75.0	76	4	US-08-981-739-19	Sequence 19, Appl
20	18	75.0	76	4	US-08-981-739-22	Sequence 22, Appl
21	18	75.0	76	1	US-08-519-777-23	Sequence 23, Appl
22	18	75.0	76	1	US-08-519-777-24	Sequence 24, Appl
23	18	75.0	95	1	US-08-742-035-23	Sequence 23, Appl
24	18	75.0	95	1	US-08-742-035-24	Sequence 24, Appl
25	18	75.0	95	2	US-08-777-019-23	Sequence 23, Appl
26	18	75.0	95	2	US-08-777-019-24	Sequence 24, Appl
27	18	75.0	95	2	US-08-777-143-23	Sequence 23, Appl
28	18	75.0	95	2	US-08-777-143-24	Sequence 24, Appl
29	18	75.0	95	3	US-08-775-414-23	Sequence 23, Appl
30	18	75.0	95	3	US-08-775-414-24	Sequence 24, Appl
31	18	75.0	95	4	US-08-931-858E-23	Sequence 23, Appl
32	18	75.0	95	4	US-08-931-858E-24	Sequence 24, Appl
33	18	75.0	95	4	US-08-981-739-23	Sequence 23, Appl
34	18	75.0	119	4	US-08-981-739-24	Sequence 24, Appl
35	18	75.0	95	4	US-09-046-479-4	Sequence 4, Appli
36	18	75.0	119	1	US-08-420-235B-23	Sequence 23, Appl
37	18	75.0	121	4	US-08-793-624-23	Sequence 23, Appl
38	18	75.0	121	5	PCT-US95-10194-23	Sequence 23, Appl
39	18	75.0	121	4	US-08-159-784-23	Sequence 23, Appl
40	18	75.0	142	3	US-08-775-414-82	Sequence 82, Appli
41	18	75.0	144	3	US-08-775-414-81	Sequence 81, Appli
42	18	75.0	150	3	US-08-775-414-84	Sequence 84, Appli
43	18	75.0	152	3	US-08-775-414-83	Sequence 83, Appli
44	18	75.0	191	1	US-08-159-784-23	Sequence 23, Appli
45	18	75.0	195	1	US-08-519-777-8	Sequence 8, Appli
46	18	75.0	195	2	US-08-742-035-8	Sequence 8, Appli
47	18	75.0	195	2	US-08-777-019-8	Sequence 8, Appli
48	18	75.0	195	3	US-08-777-143-8	Sequence 8, Appli
49	18	75.0	195	4	US-08-777-019-24	Sequence 24, Appli
50	18	75.0	195	2	US-08-852-401-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-08-852-401-3
Sequence 3, Application US/08852401
Patent No. 5976836
GENERAL INFORMATION:
  APPLICANT: Weber, J. Mark
  APPLICANT: Hessler, Paul E.
  APPLICANT: Larsen, Peter E.
  APPLICANT: Luu, B. Minh
  TITLE OF INVENTION: Methods and Compositions for Enhancing
  TITLE OF INVENTION: Erythromycin Production
  NUMBER OF SEQUENCES: 5
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Rockey, Milnamow & Katz, Ltd.
    STREET: 2 Prudential Plaza, 180 N. Stetson, Suite
    STREET: 4700
    CITY: Chicago
    STATE: Illinois
    COUNTRY: U.S.A.
    ZIP: 60601
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.30
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/852,401
    FILING DATE:
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Mueller, Lisa L.
    REGISTRATION NUMBER: 38,978
    REFERENCE/DOCKET NUMBER: FER2159PO03OUS
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 312-616-5400
    TELEFAX: 312-616-5460
INFORMATION FOR SEQ ID NO: 3:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 526 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
  MOLECULE TYPE: peptide

Query Match
Best Local Similarity 79.2%;   Score 19; DB 2;   Length 526;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   1 LXXYRXXL 8
       | ||| |
Db 447 LSTYRKAL 454
```

RESULT 2
US-08-537-002A-3
Sequence 3, Application US/08537002A
Patent No. 5773282
GENERAL INFORMATION:
APPLICANT: TSUSAKI, Keiji
APPLICANT: KUBOTA, Michio
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME FOR
TITLE OF INVENTION: CONVERTING MALTOSE INTO TREHALOSE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/537,002A
FILING DATE: 29-SEP-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 260084/1994
FILING DATE: 01-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP No. 5773282 yet received
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 963 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-537-002A-3

Query Match
Best Local Similarity 79.2%;   Score 19; DB 1;   Length 963;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   1 LXXYRXXL 8
       | ||| |
Db 912 LQAYRSAL 919
```

RESULT 3
US-08-863-010-3
; Sequence 3, Application US/08863010

Patent No. 6087146
GENERAL INFORMATION:
APPLICANT: TSUSAKI, Keiji
APPLICANT: KUBOTA, Michio
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME FOR
TITLE OF INVENTION: CONVERTING MALTOSE INTO TREHALOSE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/863,010
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/537,002
FILING DATE: 29-SEP-1995
APPLICATION NUMBER: JP 260084/1994
FILING DATE: 01-OCT-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP No. 6087146 yet received
FILING DATE: 08-SEP-1995
ATTORNEY/AGENT INFORMATION:
NAME: BROWDY, Roger L.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TSUSAKI-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 963 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-863-010-3

Query Match
Best Local Similarity 79.2%;   Score 19; DB 3;   Length 963;
Matches 4; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

```
Qy   1 LXXYRXXL 8
       | ||| |
Db 912 LQAYRSAL 919
```

RESULT 4
US-09-024-429-3
; Sequence 3, Application US/09024429
Patent No. 6165768
GENERAL INFORMATION:
APPLICANT: TSUSAKI, Keiji
APPLICANT: KUBOTA, Michio
APPLICANT: SUGIMOTO, Toshiyuki
TITLE OF INVENTION: RECOMBINANT THERMOSTABLE ENZYME FOR
TITLE OF INVENTION: CONVERTING MALTOSE INTO TREHALOSE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 419 Seventh Street, N.W., Suite 300

```
                                GenCore version 4.5
                                Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:          April 21, 2002, 10:17:58 ; Search time 102.84 Seconds
                                             (without alignments)
                                             5.185 Million cell updates/sec

Title:           US-09-281-760C-4
Perfect score:   35
Sequence:        1 TLLEYRM 7
Scoring table:   BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched:        219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters:   219241

Database :       PIR.68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

                 Minimum DB seq length: 0
                 Maximum DB seq length: 200000000

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	85.7	171	2	C86750	hypothetical prote
2	30	85.7	308	2	H65035	hypothetical prote
3	30	85.7	308	2	A85904	probable transcrip
4	30	85.7	324	2	E70532	hypothetical prote
5	30	85.7	326	2	S72945	hypothetical prote
6	30	85.7	421	2	S38110	O-sialoglycoprotei
7	30	85.7	2368	2	S46005	ESR1 protein - yea
8	29	82.9	267	2	T33911	hypothetical prote
9	29	82.9	315	2	T33910	hypothetical prote
10	29	82.9	344	2	T33909	hypothetical prote
11	29	82.9	388	2	T33908	hypothetical prote
12	29	82.9	531	2	T21059	hypothetical prote
13	29	82.9	631	2	C96768	hypothetical prote
14	29	82.9	988	2	A40628	probable transposa
15	29	82.9	2150	2	T08165	RNA1 polyprotein -
16	28	80.0	223	2	S48765	probable membrane
17	28	80.0	256	2	A69129	conserved hypothet
18	28	80.0	501	2	A29487	cytochrome P450 3A
19	28	80.0	501	2	A34236	cytochrome P450 3A
20	28	80.0	526	2	P5XRBU	outer capsid prote
21	28	80.0	526	1	P5XX15	outer capsid prote
22	28	80.0	526	2	S18768	outer capsid prote
23	28	80.0	617	2	T23952	outer capsid prote
24	28	80.0	851	2	T41041	outer capsid prote
25	28	80.0	919	2	B72765	outer capsid prote
26	28	80.0	617	2	G70685	probable nucleopor
27	27	77.1	126	2	C70865	hypothetical prote
28	27	77.1	214	2	D70813	hypothetical prote
29	27	77.1	216	2	D70813	endopeptidase Clp
30	27	77.1	246	1	G64166	heme export protei
31	27	77.1	265	2	T20127	hypothetical prote
32	27	77.1	353	2	A41558	N-syndecan - rat (
33	27	77.1	384	2	F82030	probable GTP-bindi
34	27	77.1	384	2	D81007	GTP-binding protei
35	27	77.1	386	2	F82721	two-component syst
36	27	77.1	576	2	A84898	conserved hypothet
37	27	77.1	594	2	D71347	conserved hypothet
38	27	77.1	653	2	T47581	SCARECROW1 - Arabi
39	27	77.1	653	2	T51244	SCARECROW protein
40	27	77.1	783	2	T38690	probable regulator
41	27	77.1	1160	2	T00272	hypothetical prote
42	26	74.3	82	2	I50968	hypothetical prote
43	26	74.3	82	2	I50984	MHC class II beta
44	26	74.3	82	2	I50986	MHC class II beta
45	26	74.3	237	2	T10629	MHC class I beta
46	26	74.3	255	2	S74068	hypothetical prote
47	26	74.3	282	2	JC7333	multicystatin - co
48	26	74.3	290	2	S50050	hypothetical prote
49	26	74.3	313	2	T23675	hypothetical prote
50	26	74.3	318	2	T02288	hypothetical prote

ALIGNMENTS

RESULT 1
C86750
hypothetical protein ykdA [imported] - Lactococcus lactis subsp. lactis (strain IL140
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 23-Mar-2001
C;Accession: C86750
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarme, K.; Weissenbach, J.; Eh
Genome Res. in press, 2001
A;Title:The complete genome sequence of the lactic acid bacterium.
A;Reference number: A86625
A;Accession: C86750
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-171 <STO>
A;Cross-references: GB:AE005176; NID:g12723950; PIDN:AAK05101.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: ykdA

```
Qy     1 TLLEYR 6
           ||||||
Db   145 TLLEYR 150
```

```
Query Match            85.7%;   Score 30;  DB 2;  Length 171;
Best Local Similarity  100.0%;  Pred. No. 13;
Matches     6;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;
```

RESULT 2
H65035
hypothetical protein srmB-ung intergenic region - Escherichia coli (strain K-12)
C;Species: Escherichia coli
C;Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 17-Mar-2000
C;Accession: H65035
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A;Title:The complete genome sequence of Escherichia coli K-12.
A;Reference number: A64720; MUID:97426617
A;Accession: H65035
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-308 <BLAT>
A;Cross-references: GB:U00096; NID:g1788927; PIDN:AAC75630.1; PID:g17889
A;Experimental source: strain K-12, substrain MG1655
C;Genetics:

A;Gene: yfiE
C;Superfamily: Pseudomonas putida regulatory protein catR

Query Match   85.7%; Score 30; DB 2; Length 308;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 TLLEYRM 7
          ||| |||
Db 114 TLLSYRM 120

RESULT 3
A85904
probable transcription regulator LYSR-type yfiE [imported] - Escherichia coli (strain O1...
C;Species: Escherichia coli
C;Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 31-Mar-2001
C;Accession: A85904
R;Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A;Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A;Reference number: A85904
A;Accession: A85904
A;Accession: A85480; MUID:21074935; PMID:11206551
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-308 <STO>
A;Cross-references: GB:AE005174; NID:g12516998; PIDN:AAG57693.1; GSPDB:GN00145; UWGP:238
A;Genetics:
A;Experimental source: strain O157:H7, substrain EDL933
A;Gene: yfiE

Query Match   85.7%; Score 30; DB 2; Length 308;
Best Local Similarity 85.7%; Pred. No. 25;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy   1 TLLEYRM 7
          ||| |||
Db 114 TLLSYRM 120

RESULT 4
E70532
hypothetical protein Rv2714 - Mycobacterium tuberculosis (strain H37RV)
C;Species: Mycobacterium tuberculosis
C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C;Accession: E70532
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A;Authors: Squres, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A;Reference number: E70532
A;Accession: E70532
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-324 <COL>
A;Cross-references: GB:Z96072; GB:AL123456; NID:g3261793; PIDN:CAB09474.1; PID:e320946;
A;Genetics:
A;Experimental source: strain H37Rv
A;Gene: Rv2714

Query Match   85.7%; Score 30; DB 2; Length 324;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   2 LLEYRM 7
          ||||||
Db 194 LLEYRM 199

RESULT 5
S72945
hypothetical protein B2235_F1_6 - Mycobacterium leprae
C;Species: Mycobacterium leprae
C;Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 23-Mar-2001
C;Accession: S72945
R;Smith, D.R.; Robison, K.
submitted to the EMBL Data Library, November 1993
A;Description: Mycobacterium leprae cosmid B2235.
A;Reference number: S72587
A;Accession: S72945
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-326 <SMI>
A;Cross-references: EMBL:U00019; NID:g467079; PIDN:AAA17281.1; PID:g467098

Query Match   85.7%; Score 30; DB 2; Length 326;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   2 LLEYRM 7
          ||||||
Db 196 LLEYRM 201

RESULT 6
S38110
O-sialoglycoprotein endopeptidase homolog YKR038c - yeast (Saccharomyces cerevisiae)
O;sialoglycoprotein endopeptidase
C;Superfamily: O-sialoglycoprotein endopeptidase
C;Species: Saccharomyces cerevisiae
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 11-Jan-2000
C;Accession: S38110
R;Urrestarazu, L.A.; Jauniaux, J.C.
submitted to the Protein Sequence Database, March 1994
A;Reference number: S38097
A;Accession: S38110
A;Molecule type: DNA
A;Residues: 1-421 <URR>
A;Cross-references: EMBL:Z28263; NID:g486476; PID:g486477; MIPS:YKR038c
A;Experimental source: strain S288C
A;Genetics:
A;Map position: 11R
C;Keywords: calcium binding; EF hand
F;271-283/Domain: calcium binding #status predicted <CAl>

Query Match   85.7%; Score 30; DB 2; Length 421;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   2 LLEYRM 7
          ||||||
Db 389 LLEYRM 394

RESULT 7
S46005
ESR1 protein - yeast (Saccharomyces cerevisiae)
N;Alternate names: MEC1 protein; protein YBR1012; protein YBR136w
C;Species: Saccharomyces cerevisiae
C;Date: 26-Aug-1994 #sequence_revision 09-Sep-1994 #text_change 29-Oct-1999
C;Accession: S46005
C;Accession: S46005; S46578; S46663; S47954; S64650
R;Becam, A.M.; Herbert, C.J.; Nasr, F.; Slonimski, P.P.; Zagulski, M.
submitted to the Protein Sequence Database, August 1994
A;Reference number: S45995
A;Accession: S46005
A;Molecule type: DNA
A;Residues: 1-2368 <BEC>
A;Cross-references: EMBL:Z36005; NID:g536429; PIDN:CAA85094.1; PID:g536430; MIPS:YBR1
A;Experimental source: strain S288C

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:37:25 ; Search time 85.5 Seconds
                                       (without alignments)
                                       3.002 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7
Scoring table: BLOSUM62
Database :   Gapop 10.0 , Gapext 0.5
                 SwissProt_39.*

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

imum DB seq length: 0
imum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

   Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES.

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	85.7	293	1	YFIE_ECOLI	P33634 escherichia
2	30	85.7	421	1	YK18_YEAST	P36132 saccharomyc
3	30	85.7	2368	1	ESR1_YEAST	P38111 saccharomyc
4	29	82.9	589	1	LU1A_LYCPN	O04973 lycopersico
5	29	82.9	612	1	LU1B_LYCPN	O00974 lycopersico
6	29	82.9	988	1	TNP6_ENTFC	O06238 enterococcu
7	28	80.0	223	1	YD78_YEAST	P38957 saccharomyc
8	28	80.0	432	1	GBA1_CRYNE	P54853 cryptococcu
9	28	80.0	501	1	CP36_RABIT	P11707 oryctolagus
10	28	80.0	526	1	VP5_BTV1A	P12436 bluetongue
11	28	80.0	526	1	VP5_BTV1S	P33475 bluetongue
12	28	80.0	526	1	VP5_BTV2A	P33029 bluetongue
13	28	80.0	967	1	Y5GO_CLOAB	P33747 clostridium
14	27	77.1	214	1	CL2_MYCLE	Q9cby4 mycobacteri
15	27	77.1	214	1	CL2_MYCTU	P45034 mycobacteri
16	27	77.1	246	1	CCMC_HAEIN	O75056 haemophilus
17	27	77.1	384	1	SDC3_HUMAN	Q6451 homo sapien
18	27	77.1	433	1	Y6B1_PASMU	P57864 pasteurella
19	27	77.1	442	1	SDC3_MOUSE	Q33671 mus musculu
20	27	77.1	442	1	SDC3_RAT	Q9y4e6 rattus norv
21	27	77.1	1160	1	WDR7_HUMAN	O24689 homo sapien
22	26	74.3	213	1	RL3_SYNP6	Q9z723 synechococc
23	26	74.3	436	1	CG2B_DICDI	P42524 dictyosteli
24	26	74.3	505	1	SYE_CHLPN	P05494 chlamydia p
25	26	74.3	508	1	ATPO_MAIZE	P24459 zea mays (m
26	26	74.3	508	1	ATPO_PHAVU	P15998 phaseolus v
27	26	74.3	509	1	ATPO_ORYSA	P12862 oryza sativ
28	26	74.3	509	1	ATPO_WHEAT	O03611 triticum ae
29	26	74.3	741	1	TAX4_CAEEL	P02459 caenorhabdi
30	26	74.3	747	1	CA12_BOVIN	P44403 bos taurus
31	26	74.3	856	1	CLPB_HAEIN	O06327 haemophilus
32	26	74.3	859	1	LOX1_ARATH	Q09731 arabidopsis
33	26	74.3	962	1	YA4E_SCHPO	schizosach
34	26	74.3	1362	1	CA21_CHICK	P02467 gallus gall
35	26	74.3	1595	1	SOS_DROME	P26675 drosophila
36	25	71.4	29	1	HS98_NEUCR	P31540 neurospora
37	25	71.4	82	1	VPU_HV1BN	P12516 human immun
38	25	71.4	111	1	RL29_MYCPN	Q50310 mycoplasma
39	25	71.4	153	1	RL29_MYCGA	O52240 mycoplasma
40	25	71.4	183	1	IMP3_YEAST	P32899 saccharomyc
41	25	71.4	195	1	YI88_BACHD	Q9kbn7 bacillus ha
42	25	71.4	200	1	RL29_MYCGE	P47405 mycoplasma
43	25	71.4	208	1	EFH1_TRYCR	P41048 trypanosoma
44	25	71.4	208	1	EFH2_TRYCR	P41049 trypanosoma
45	25	71.4	211	1	YL71_ARCFU	O28111 archaeoglob
46	25	71.4	214	1	YO23_CAEEL	O34673 caenorhabdi
47	25	71.4	226	1	ATP6_BALMU	P41291 balaenopter
48	25	71.4	227	1	MTR2_MOUSE	Q9z2d1 mus musculu
49	25	71.4	234	1	AOX2_AERPE	Q9ydx7 aeropyrum p
50	25	71.4	238	1	VSP1_AGKHA	P81176 agkistrodon

ALIGNMENTS

RESULT 1
YFIE_ECOLI
ID   YFIE_ECOLI   STANDARD;   PRT;   293 AA.
AC   P33634; P76592; P76998;
DT   01-FEB-1994 (Rel. 28, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   HYPOTHETICAL TRANSCRIPTIONAL REGULATOR YFIE.
GN   YFIE OR B2577.
OS   Escherichia coli.
OC   Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC   Escherichia.
OX   NCBI_TaxID=562;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RA   Nashimoto H.;
RT   "Non-ribosomal proteins affecting the assembly of ribosomes in
RT   Escherichia coli.";
RL   (In) Nierhaus K.H. (eds.);
RL   The translational apparatus, pp.185-195, Plenum Press,
RL   New York (1993).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12;
RA   Nashimoto H., Saito N.;
RL   Submitted (SEP-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=K12 / MG1655;
RX   MEDLINE=97426617; PubMed=9278503;
RA   Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA   Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA   Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA   Mau B., Shao Y.;
RT   "The complete genome sequence of Escherichia coli K-12.";
RL   Science 277:1453-1474(1997).
RN   [4]
RP   SEQUENCE OF 1-208 FROM N.A.
RC   STRAIN=K12;
RX   MEDLINE=97349980; PubMed=9205837;
RA   Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
RA   Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
RA   Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
RA   Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasundaram S.,
RA   Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
RA   Yamagata S., Yamada M., Yano M., Yasuda S., Horiuchi T.;
RT   "Construction of a contiguous 874-kb sequence of the Escherichia coli
RT   - K12 genome corresponding to 50.0-68.8 min on the linkage map and
RT   analysis of its sequence features.";
RL   DNA Res. 4:91-113(1997).

```
CC   -1-  SIMILARITY: BELONGS TO THE LYSR FAMILY OF TRANSCRIPTIONAL
CC        REGULATORS.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; D13169; -; NOT_ANNOTATED_CDS.
DR   EMBL; D64044; -; NOT_ANNOTATED_CDS.
DR   EMBL; AE000344; AAC75630.1; ALT_INIT.
DR   EMBL; D90886; BAA20920.1; -.
DR   EcoGene; EG11785; yfiE.
DR   InterPro; IPR000847; HTH_LysR.
DR   Pfam; PF00126; HTH_1; 1.
KW   PROSITE; PS00044; HTH_LYSR_FAMILY; 1.
KW   Hypothetical protein; Transcription regulation; DNA-binding;
KW   Complete proteome.
DNA_BIND     18     37   H-T-H MOTIF (BY SIMILARITY).
CONFLICT    209    293

SEQUENCE   293 AA;  33243 MW;  1C54C96647799632 CRC64;
     RITVENTIELISTESIKRCVAANIGVSYLPREAVAKELECG
     ELIELPFGEQSQITTAMCAHHAGKAVSPAMHFIQCVEESF
     VAG -> DHGGKHHI (IN REF. 1 AND 2).


Qy    1  TLLEYRM  7
         |||||||
Db   99  TLLSYRM  105

Query Match             85.7%;  Score 30;  DB 1;  Length 293;
Best Local Similarity   85.7%;  Pred. No. 11;
Matches  6;  Conservative  0;  Mismatches  1;  Indels  0;  Gaps  0;


RESULT 2
ID   YK18_YEAST     STANDARD;      PRT;     421 AA.
AC   P36132;
DT   01-JUN-1994 (Rel. 29, Created)
DT   01-JUN-1994 (Rel. 29, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   HYPOTHETICAL 46.6 KDA PROTEIN IN DAL80-GAP1 INTERGENIC REGION.
GN   YKR038C.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Urrestarazu L.A., Jauniaux J.-C.;
RL   Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: COULD BE A METALLOPROTEASE.
CC   -!- COFACTOR: ZINC (PROBABLE).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M22; ALSO KNOWN AS THE
CC       GLYCOPROTEASE FAMILY.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; Z28263; CAA82112.1; -.
DR   PIR; S38110; S38110.
DR   SGD; S0001746; YKR038C.
DR   InterPro; IPR000905; Peptidase_M22.
DR   Pfam; PF00814; Peptidase_M22; 1.
DR   PRINTS; PR00789; OSIALOPTASE.
DR   ProDom; PD003367; Peptidase_M22; 1.
DR   PROSITE; PS01016; GLYCOPROTEASE; 1.
KW   Hypothetical protein; Hydrolase; Metalloprotease; Zinc.
FT   METAL    176    176   ZINC (POTENTIAL).
FT   METAL    180    180   ZINC (POTENTIAL).
SQ   SEQUENCE   421 AA;  46602 MW;  0572114980222772F CRC64;


Qy    2  LLEYRM  7
         ||||||
Db  389  LLEYRM  394

Query Match             85.7%;  Score 30;  DB 1;  Length 421;
Best Local Similarity  100.0%;  Pred. No. 16;
Matches  6;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT 3
ID   ESR1_YEAST     STANDARD;      PRT;    2368 AA.
AC   P38111; Q02580;
DT   01-OCT-1994 (Rel. 30, Created)
DT   01-OCT-1994 (Rel. 30, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   ESR1 PROTEIN.
GN   ESR1 OR MEC1 OR SAD3 OR YBR136W OR YBR1012.
OS   Saccharomyces cerevisiae (Baker's yeast).
OC   Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC   Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX   NCBI_TaxID=4932;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94344772; PubMed=8065923;
RA   Kato R., Ogawa H.;
RT   "An essential gene, ESR1, is required for mitotic cell growth, DNA
RT   repair and meiotic recombination in Saccharomyces cerevisiae.";
RL   Nucleic Acids Res. 22:3104-3112(1994).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Weinert T.A., Harlow D.;
RL   Submitted (JUL-1995) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=S288C;
RX   MEDLINE=94378717; PubMed=8091856;
RA   Becam A.M., Cullin C., Grzybowska E., Lacroute F., Nasr F.,
RA   Ozier-Kalogeropoulos O., Palucha A., Slonimski P.P., Zagulski M.,
RA   Herbert C.J.;
RT   "The sequence of 29.7 kb from the right arm of chromosome II reveals
RT   13 complete open reading frames, of which ten correspond to new
RT   genes.";
RL   Yeast 10:511-511(1994).
CC   -!- FUNCTION: REQUIRED FOR CELL GROWTH, DNA REPAIR AND MEIOTIC
CC       RECOMBINATION.
CC   -!- DEVELOPMENTAL STAGE: INDUCED DURING MEIOSIS.
CC   -!- SIMILARITY: IN THE N-TERMINAL REGION, TO S.POMBE CUT1.
CC   -!- SIMILARITY: BELONGS TO THE PI3/PI4-KINASES FAMILY.
CC   --------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   --------------------------------------------------------------
DR   EMBL; X75891; CAA53494.1; -.
DR   EMBL; U31109; AAA74482.1; -.
DR   EMBL; Z36005; CAA85094.1; -.
DR   EMBL; D11088; BAA01860.1; -.
DR   PIR; S46005; S46005.
DR   PIR; S46578; S46578.
DR   SGD; S0000340; MEC1.
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:35:50 ; Search time 191.29 Seconds
(without alignments)
5.353 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEVRM 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues
Listing first 50 summaries
Total number of hits satisfying chosen parameters: 473505
Maximum DB seq length: 200000000
Minimum DB seq length: 0
Post-processing: Minimum Match 0%
Maximum Match 100%

Database :
```
SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	85.7	171	2	Q9JN51	lactococcus
2	30	85.7	171	2	Q9CGT9	lactococcus
3	30	85.7	324	2	O07213	mycobacteri
4	30	85.7	326	2	O49847	mycobacteri
5	30	85.7	330	2	Q9CPB9	pasteurella
6	30	85.7	422	2	Q9LCY8	thermus aqu
7	29	82.9	267	5	Q9TYL5	caenorhabdi
8	29	82.9	315	5	Q9TYL6	caenorhabdi
9	29	82.9	344	5	Q9TYL7	caenorhabdi
10	29	82.9	388	5	Q9TYL8	caenorhabdi
11	29	82.9	531	5	O19518	caenorhabdi
12	29	82.9	631	10	Q9C550	arabidopsis
13	29	82.9	675	12	Q9LPR4	arabidopsis
14	29	82.9	2150	1	O37086	peach roset
15	28	80.0	256	1	O26335	methanobact
16	28	80.0	565	2	Q9REN5	zymomonas m
17	28	80.0	587	4	O95876	homo sapien
18	28	80.0	617	5	O9U398	caenorhabdi
19	28	80.0	851	3	O94418	schizosacch
20	28	80.0	919	1	Q9YFZ1	aeropyrum p
21	27	77.1	126	2	O86327	mycobacteri
22	27	77.1	233	2	O53856	mycobacteri
23	27	77.1	233	2	Q9RP41	cytophaga j
24	27	77.1	247	10	O41171	rhodella vi
25	27	77.1	265	5	O18763	caenorhabdi
26	27	77.1	359	13	Q9IBF8	oryzias lat
27	27	77.1	378	2	Q9L433	xanthomonas
28	27	77.1	382	13	Q9DG99	oryzias lat
29	27	77.1	384	2	Q9JXE5	neisseria m
30	27	77.1	384	2	Q9JW74	neisseria m
31	27	77.1	386	2	Q9PEB5	neisseria m
32	27	77.1	386	13	Q9DG96	oryzias luz
33	27	77.1	387	13	Q91BG0	oryzias lat
34	27	77.1	388	13	Q9DGA3	oryzias cur
35	27	77.1	409	8	Q9MR08	euglypha ro
36	27	77.1	419	8	Q9RM06	helicobacte
37	27	77.1	453	5	Q9NKF8	drosophila
38	27	77.1	481	5	Q9VTW8	drosophila
39	27	77.1	530	5	Q9GPF0	dictyosteli
40	27	77.1	576	5	Q9W329	drosophila
41	27	77.1	576	10	O82364	arabidopsis
42	27	77.1	594	2	O83287	treponema p
43	27	77.1	653	10	Q96304	arabidopsis
44	27	77.1	653	10	Q9M384	arabidopsis
45	27	77.1	762	5	Q9U2F2	caenorhabdi
46	27	77.1	774	4	P95636	rhodopseudo
47	27	77.1	783	3	O14130	schizosacch
48	27	77.1	853	4	Q9BS12	homo sapien
49	27	77.1	1024	2	Q9X08	streptococc
50	27	77.1	1488	11	Q9ERH3	rattus norv

ALIGNMENTS

RESULT 1
Q9JN51
```
ID   Q9JN51      PRELIMINARY;      PRT;   171 AA.
AC   Q9JN51;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   TCSEORF2.
OS   Lactococcus lactis (subsp. cremoris) (Streptococcus cremoris).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC   Lactococcus.
OX   NCBI_TaxID=1359;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MG1363;
RX   MEDLINE=97316457; PubMed=9172368;
RA   O'Connell-Motherway M., Fitzgerald G.F., van Sinderen D.;
RT   "Cloning and sequence analysis of putative histidine protein kinases
RT   isolated from Lactococcus lactis MG1363.";
RL   Appl. Environ. Microbiol. 63:2454-2459(1997).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=MG1363;
RX   MEDLINE=20244638; PubMed=10784052;
RA   O'Connell-Motherway M., van Sinderen D., Morel-Deville F.,
RA   Fitzgerald G.F., Ehrlich S.D., Morel P.;
RT   "Six putative two-component regulatory systems isolated from
RT   Lactococcus lactis subsp. cremoris MG1363.";
RL   Microbiology 146:935-947(2000).
DR   EMBL; U81488; AAF66143.1; -.
SQ   SEQUENCE   171 AA;  19171 MW;  3D16FAD7BFF83819 CRC64;
```

Query Match 85.7%; Score 30; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy         1 TLLEYR 6
             ||||||
Db       145 TLLEYR 150
```

```
RESULT   2
ID       Q9CGT9
AC       Q9CGT9;
DT       01-JUN-2001 (TrEMBLrel. 17, Created)          PRELIMINARY;      PRT;    171 AA.
DT       01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT       01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE       UNKNOWN PROTEIN.
GN       YKDA.
OS       Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC       Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC       Lactococcus.
OX       NCBI_TaxID=1360;
RN       [1]
RP       SEQUENCE FROM N.A.
RC       STRAIN=IL1403;
RA       Bolotin A., Wincker P., Mauger S., Jaillon O., Malarme K.,
RA       Weissenbach J., Ehrlich S.D., Sorokin A.;
RT       "The complete genome sequence of the lactic acid bacterium Lactococcus
RT       lactis.";
RL       Genome Res. 0:0-0(2001).
DR       EMBL; AE006334; AAK05101.1; -.
KW       Complete proteome.
SQ       SEQUENCE  171 AA;   19216 MW;  975EEFBC50329BF9 CRC64;
```

```
Qy         1 TLLEYR 6
             ||||||
Db       145 TLLEYR 150
```

```
Query Match             85.7%;  Score 30;  DB 2;  Length 171;
Best Local Similarity  100.0%;  Pred. No. 28;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;
```

```
Qy         2 LLEYRM 7
             ||||||
Db       194 LLEYRM 199
```

```
RESULT   3
ID       O07213
AC       O07213;
DT       01-JUL-1997 (TrEMBLrel. 04, Created)          PRELIMINARY;      PRT;    324 AA.
DT       01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT       01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE       HYPOTHETICAL 35.5 KDA PROTEIN.
GN       RV2714 OR MTCY05A6.35.
OS       Mycobacterium tuberculosis.
OC       Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC       Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX       NCBI_TaxID=1773;
RN       [1]
RP       SEQUENCE FROM N.A.
RC       STRAIN=H37RV;
RX       MEDLINE=98295587; PubMed=9634230;
RA       Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA       Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaia F.,
RA       Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA       Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA       Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA       Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA       Rutter S., Seeger K., Skelton S., Squares S., Squares R., Sulston J.E.,
RA       Taylor K., Whitehead S., Barrell B.G.;
RT       "Deciphering the biology of Mycobacterium tuberculosis from the
RT       complete genome sequence.";
RL       Nature 393:537-544(1998).
DR       EMBL; Z96072; CAB09474.1; -.
KW       Tuberculosis; Hypothetical protein; Complete proteome.
SQ       SEQUENCE  324 AA;   35519 MW;  45E48809917AB22B CRC64;
```

```
Qy         2 LLEYRM 7
             ||||||
Db       196 LLEYRM 201
```

```
RESULT   4
ID       Q49847
AC       Q49847;
DT       01-NOV-1996 (TrEMBLrel. 01, Created)          PRELIMINARY;      PRT;    326 AA.
DT       01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT       01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE       B2235_F1.6 (HYPOTHETICAL PROTEIN ML1009).
GN       ML1009.
OS       Mycobacterium leprae.
OC       Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC       Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX       NCBI_TaxID=1769;
RN       [1]
RP       SEQUENCE FROM N.A.
RA       Robison K.;
RL       Submitted (NOV-1993) to the EMBL/GenBank/DDBJ databases.
RN       [2]
RP       SEQUENCE FROM N.A.
RA       Robison K.;
RL       Submitted (MAR-1994) to the EMBL/GenBank/DDBJ databases.
RN       [3]
RP       SEQUENCE FROM N.A.
RA       Smith D.R.;
RL       Submitted (JAN-1994) to the EMBL/GenBank/DDBJ databases.
RN       [4]
RP       SEQUENCE FROM N.A.
RC       STRAIN=TN;
RX       MEDLINE=21128732; PubMed=11234002;
RA       Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
RA       Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA       Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA       Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA       Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA       Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA       Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA       Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA       Barrell B.G.;
RT       "Massive gene decay in the leprosy bacillus.";
RL       Nature 409:1007-1011(2001).
DR       EMBL; U00011; AAA17281.1; -.
DR       EMBL; U00019; AAA17281.1; -.
DR       EMBL; AL583920; CAC31390.1; -.
KW       Hypothetical protein; Complete proteome.
SQ       SEQUENCE  326 AA;   36152 MW;  D089CF228598C722 CRC64;
```

```
Qy         2 LLEYRM 7
             ||||||
Db       194 LLEYRM 199
```

```
Query Match             85.7%;  Score 30;  DB 2;  Length 324;
Best Local Similarity  100.0%;  Pred. No. 55;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;
```

```
Query Match             85.7%;  Score 30;  DB 2;  Length 326;
Best Local Similarity  100.0%;  Pred. No. 56;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Gaps      0;
```

```
RESULT   5
ID       Q9CPB9
AC       Q9CPB9;
DT       01-JUN-2001 (TrEMBLrel. 17, Created)          PRELIMINARY;      PRT;    330 AA.
DT       01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT       01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE       FECD.
```

```
OM protein - protein search, using sw model

                    Copyright (c) 1993 - 2000 Compugen Ltd.

                    GenCore version 4.5


Run on:             April 21, 2002, 10:14:10 ; Search time 158.02 Seconds
                                              (without alignments)
                                              3.281 Million cell updates/sec

Title:              US-09-281-760C-4
Perfect score:      35
Sequence:           1 TLLEYRM 7
Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters:       522463

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :      A_Geneseq_1101:*

                 1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
                 2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
                 3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
                 4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
                 5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
                 6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
                 7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
                 8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
                 9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
                10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
                11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
                12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
                13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
                14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
                15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
                16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
                17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
                18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
                19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
                20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
                21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
                22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*
```

SUMMARIES

```
Result                  %      Query
No.    Score    Match Length DB    DB ID      Description
-----  -----    ----- ------ --   -----       -----------
  1    35      100.0    7    21    AAY69584    Immunogenic peptid
  2    35      100.0    7    21    AAY50878    Antibody 8H.8 bind
  3    30       85.7    7    21    AAY59253    Yeast glycoproteas
  4    30       85.7    7    18    AAW26663    Yeast checkpoint c
  5    30       85.7   2368   20   AAW73893    Yeast MEC1 protein
  6    29       82.9   2368   20   AAY26187    Soybean 2-isopropy
  7    29       82.9    545   20   AAR24302    Transposase from g
  8    28       80.0    988   13   AAY41237    M. thermoautotroph
  9    28       80.0    135   20   AAB93395    Human protein sequ
 10    28       80.0    304   22   AAB94147    Human protein sequ
 11    27       77.1     10   20   AAY49789    Elastolytic peptid
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
12   27   77.1    11    7   AAP60980    Polypeptide which
13   27   77.1    38   21   AAB58532    Lung cancer associ
14   27   77.1    44   22   AAM18198    Peptide #4632 enco
15   27   77.1    53   21   AAB43458    Human cancer assoc
16   27   77.1    78   22   AAM37911    Peptide #11948 enc
17   27   77.1    89   22   AAM18904    Peptide #5338 enco
18   27   77.1    89   22   AAM31445    Peptide #5482 enco
19   27   77.1   653   18   AAW38178    Arabidopsis SCAREC
20   27   77.1   653   21   AAW28569    Arabidopsis SCAREC
21   27   77.1   807   22   AAM40412    Human polypeptide
22   27   77.1  1739   22   AAM42197    Human polypeptide
23   26   74.3   157   21   AAY86238    Human secreted pro
24   26   74.3   272   21   AAG10183    Arabidopsis thalia
25   26   74.3   272   21   AAG40522    Arabidopsis thalia
26   26   74.3   277   21   AAY74877    Neisseria meningit
27   26   74.3   279   21   AAG10182    Arabidopsis thalia
28   26   74.3   279   21   AAG40521    Arabidopsis thalia
29   26   74.3   292   22   AAG92082    C glutamicum prote
30   26   74.3   295   22   AAG92400    Corynebacterium gl
31   26   74.3   295   22   AAB79400    Corynebacterium gl
32   26   74.3   358   21   AAB79401    Corynebacterium gl
33   26   74.3   363   22   AAG10181    Arabidopsis thalia
34   26   74.3   363   22   AAG91536    C glutamicum prote
35   26   74.3   386   21   AAB79069    Arabidopsis thalia
36   26   74.3   516   22   AAG40520    Arabidopsis thalia
37   26   74.3   589   21   AAY35180    C. pneumoniae prot
38   26   74.3   589   22   AAB95645    Human protein sequ
39   26   74.3   841   21   AAG39711    Arabidopsis thalia
40   26   74.3   915   20   AAB18511    H2 homologue of pr
41   26   74.3   924   20   AAY34977    C. pneumoniae prot
42   26   74.3   924   22   AAB93084    Human protein sequ
43   26   74.3  1596   21   AAB95845    Human protein sequ
44   26   74.3  1596   21   AAY68821    Amino acid sequenc
45   26   74.3  1876   22   AAW38757    Phosphatidyl inosi
46   25   71.4    23   22   AAU00611    VH ligand-binding
47   25   71.4    37   21   AAB44726    Human secreted pro
48   25   71.4    67   21   AAG08015    Arabidopsis thalia
49   25   71.4    96   21   AAG08014    Arabidopsis thalia
50   25   71.4   103   22   AAB61578    Polynavirus protei
     25   71.4   123   18   AAW11784    Early onset Alzhei
```

ALIGNMENTS

```
RESULT    1
AAY69584

ID      AAY69584 standard; peptide; 7 AA.
XX
AC      AAY69584;
XX
DT      19-APR-2000  (first entry)
XX
DE      Immunogenic peptide SEQ ID NO:9, for anti-IgE autoantibody generation.
XX
KW      Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW      allergy; canine; prophylaxis; treatment.
XX
OS      Synthetic.
XX
PN      EP955311-A2.
XX
PD      10-NOV-1999.
XX
PF      09-APR-1999;  99EP-0107020.
XX
PR      09-APR-1998;  98US-0058332.
PR      30-MAR-1999;  99US-0281761.
XX
PA      (IDEX-) IDEXX LAB INC.
XX
PI      Lawton R, Mermer B,  Francoeur G;
XX
DR      WPI; 2000-118182/11.
```

```
XX
PT   New peptides, useful for generating canine autoantibodies against
PT   immunoglobulin E (IgE) to treat allergic symptoms -
XX
PS   Claim 5; Page 5; 38pp; English.
XX
CC   Sequences AAY69576-Y69586 and AAY69598-Y69599 represent peptides that may
CC   be used to induce anti-IgE (immunoglobulin E) autoantibodies in a dog.
CC   IgE is an important mediator of canine allergic responses, including Type
CC   I immediate hypersensitivity. It is estimated that up to 30% of all dogs
CC   suffer from allergies or allergy-related skin disorders, and allergic
CC   dermatitis is thought to affect between 3 and 15% of dogs. Common canine
CC   allergens include fleas, pollens, moulds and dust, with flea allergy
CC   being the most frequent. Once IgE has been released in response to
CC   allergen, the IgE binds to receptors on mast cells. The IgE molecules
CC   can then become cross-linked, which induces the mast cells to release
CC   histamine, which leads to the manifestation of allergic symptoms. The
CC   peptides of the invention correspond to epitopes of canine IgE, or
CC   conservative variations thereof, that are accessible in circulating IgE
CC   but not in IgE bound to mast cells. The autoantibodies produced on
CC   immunisation with such peptides therefore target only circulating IgE,
CC   avoiding the possibility that they may cross-link bound IgE molecules,
CC   which would exacerbate an ongoing allergic response even in the absence
CC   of allergen. Anti-IgE autoantibody induction results in an immune
CC   response that targets the IgE-expressing B cell, which affects IgE
CC   synthesis by downregulating IgE production by B cells and/or targeting
CC   the IgE-producing B cell for destruction. The peptides of the invention
CC   are therefore useful for the treatment or prophylaxis of allergic
CC   symptoms in a dog.
XX
SQ   Sequence   7 AA;

Qy          1 TLLEYRM 7
              |||||||
Db          1 tlleyrm 7

Query Match             100.0%;  Score 35;  DB 21;    Length 7;
Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

RESULT   2
ID   AAY50878 standard; peptide; 7 AA.
XX
AC   AAY50878;
XX
DT   24-FEB-2000   (first entry)
XX
DE   Antibody 8H.8 binding mimotope 1.
XX
KW   Canine; allergy; antibody 8H.8; IgE; B cell; mast cell; anti-allergy;
KW   epitope; prophylaxis; treatment; mimotope.
XX
OS   Synthetic.
XX
PN   EP957111-A2.
XX
PD   17-NOV-1999.
XX
PF   09-APR-1999;   99EP-0107035.
XX
PR   09-APR-1998;   98US-0058331.
PR   30-MAR-1999;   99US-0281760.
XX
PA   (IDEX-) IDEXX LAB INC.
XX
PI   Lawton R, Mermer B, Francoeur G;
XX
DR   WPI; 2000-040833/04.
XX
PT   Binding proteins used for treatment or prophylaxis of canine allergy -
```

```
XX
PT   Claim 9; Page 12; 30pp; English.
XX
CC   This invention describes a novel binding protein which specifically
CC   binds to native canine free or B-cell bound IgE, and which doesn't bind
CC   to IgE when the IgE is bound to mast cells. The peptide products of the
CC   invention have anti-IgE or anti-allergic activity. The antibodies bind to defined.
CC   epitopes on free or B-cell bound IgE molecules which have an important
CC   role in allergic reaction. The specific binding proteins are used to
CC   produce a pharmaceutical composition, preferably with a diluent, which
CC   can be used for prophylaxis or treatment of canine allergy.
CC   AAY50876-Y50900 represent peptide mimotopes used in the method of the
CC   invention.
XX
SQ   Sequence   7 AA;

Qy          1 TLLEYRM 7
              |||||||
Db          1 tlleyrm 7

Query Match             100.0%;  Score 35;  DB 21;    Length 7;
Best Local Similarity   100.0%;  Pred. No. 4.3e+05;
Matches    7;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

RESULT   3
ID   AAY59253 standard; Protein; 421 AA.
XX
AC   AAY59253;
XX
DT   25-APR-2000   (first entry)
XX
DE   Yeast glycoprotease-like protein.
XX
KW   Glycoprotease; glycoprotease-like protein; cell proliferation; human;
KW   immunostimulant; immunosuppressant; thrombotic agent; tissue growth;
KW   contraceptive; fertility-promoting agent; haematopoiesis; antimicrobial;
KW   anti-inflammatory; anti-tumour; gene therapy; yeast.
XX
OS   Saccharomyces cerevisiae.
XX
PN   WO967399-A1.
XX
PD   29-DEC-1999.
XX
PF   24-JUN-1999;   99WO-JP03358.
XX
PR   25-JUN-1998;   98JP-0178247.
XX
PA   (SAGA ) SAGAMI CHEM RES CENT.
PA   (PROT-) PROTEGENE INC.
XX
PI   Kato S, Kimura T;
XX
DR   WPI; 2000-147211/13.
XX
PT   A human protein with glycoprotease-like activity, used for diagnosis,
PT   treatment and prevention of e.g. autoimmune disease or cancer -
XX
PS   Disclosure; Page 35; 55pp; English.
XX
CC   The invention provides a new human protein with glycoprotease activity.
CC   The protein can be expressed by standard recombinant methodology. The
CC   glycoprotease-like protein can be used: to screen for specific-binding
CC   agents (receptors or ligands) or inhibitors or agonists of its binding
CC   interactions; to raise antibodies or elicit an immune response; as a
CC   tissue marker; as a nutrient source or supplement; as a research reagent;
CC   and as a therapeutic/biologically active agent, e.g. for regulating
CC   cell proliferation or differentiation; as immunostimulant or
CC   immunosuppressant; as thrombotic agent; for stimulation of tissue
CC   growth; as contraceptive/fertility-promoting agent; in regulation of
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:15:57 ; Search time 94.43 Seconds
(without alignments)
1.668 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 212252
1mun DB seq length: 0
1mun DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	85.7	2368	1	US-08-198-446B-15	Sequence 15, Appl
2	30	85.7	2368	2	US-08-870-693-15	Sequence 15, Appl
3	29	82.9	988	2	US-08-286-819A-19	Sequence 19, Appl
4	29	82.9	988	3	US-08-980-357-19	Sequence 19, Appl
5	28	80.0	135	3	US-09-053-197A-32	Sequence 32, Appl
6	27	77.1	1023	1	US-08-198-446B-4	Sequence 4, Appl
7	27	77.1	1023	2	US-08-870-693-4	Sequence 4, Appli
8	26	74.3	137	2	US-08-609-049A-14	Sequence 14, Appl
9	26	74.3	137	4	US-09-170-996-14	Sequence 14, Appl
10	26	74.3	430	2	US-08-290-731C-9	Sequence 9, Appli
11	26	74.3	1572	2	US-08-290-731C-5	Sequence 5, Appli
12	26	74.3	1596	2	US-09-356-952-3	Sequence 3, Appli
13	26	74.3	1876	2	US-08-609-049A-12	Sequence 12, Appl
14	26	74.3	1876	2	US-08-609-049A-28	Sequence 28, Appl
15	26	74.3	1876	4	US-09-170-996-12	Sequence 12, Appl
16	26	74.3	1876	4	US-09-170-996-28	Sequence 28, Appl
17	25	74.3	302	2	US-09-031-485-38	Sequence 38, Appl
18	25	74.3	302	2	US-08-847-429A-38	Sequence 38, Appl
19	25	74.3	302	3	US-09-065-474-38	Sequence 38, Appl
20	25	74.3	567	3	US-08-813-150-4	Sequence 4, Appl
21	25	74.3	690	2	US-08-619-554-8	Sequence 8, Appli
22	25	74.3	885	1	US-08-372-892-4	Sequence 4, Appli
23	25	71.4	302	1	US-08-445-640-35	Sequence 35, Appli
24	25	71.4	888	3	US-08-170-558-35	Sequence 35, Appl
25	25	71.4	888	3	US-08-447-314-35	Sequence 35, Appl
26	25	71.4	888	3	US-08-445-461-35	Sequence 35, Appl
27	25	71.4	892	1	US-07-977-434-12	Sequence 12, Appl
28	25	71.4	892	1	US-08-458-819-12	Sequence 12, Appl
29	25	71.4	892	5	PCT-US91-07035-12	Sequence 12, Appli
30	25	71.4	894	1	US-08-372-892-2	Sequence 2, Appl
31	25	71.4	894	1	US-08-445-640-34	Sequence 34, Appl
32	25	71.4	894	3	US-08-170-558-34	Sequence 34, Appl
33	25	71.4	894	3	US-08-447-314-34	Sequence 34, Appli
34	25	71.4	894	3	US-08-445-461-34	Sequence 34, Appl
35	25	71.4	1088	4	US-09-082-059-2	Sequence 2, Appl
36	25	71.4	894	4	US-08-619-198-3	Sequence 3, Appl
37	25	71.4	1389	2	US-08-619-198-5	Sequence 5, Appl
38	25	71.4	1418	3	US-08-963-825-20	Sequence 20, Appl
39	25	71.4	1418	4	US-00-010-999-1	Sequence 1, Appli
40	25	71.4	1442	2	US-08-316-650-12	Sequence 12, Appl
41	25	71.4	1442	5	PCT-US95-02251-12	Sequence 12, Appli
42	25	71.4	1876	2	US-08-619-554-4	Sequence 4, Appli
43	25	71.4	1895	4	US-07-977-630-48	Sequence 48, Appl
44	24	68.6	28	1	US-09-053-197A-8	Sequence 8, Appli
45	24	68.6	264	3	US-08-915-498B-27	Sequence 27, Appl
46	24	68.6	346	3	US-08-372-652-2	Sequence 2, Appl
47	24	68.6	451	2	US-08-372-183-2	Sequence 2, Appli
48	24	68.6	451	2	PCT-US95-16311-2	Sequence 2, Appli
49	24	68.6	451	5	US-09-469-721-2	Sequence 2, Appli
50	24	68.6	469	4	US-09-469-721-2	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-198-446B-15
Sequence 15, Application US/08198446B
Patent No. 5674996
GENERAL INFORMATION:
APPLICANT: Hartwell, Leland H.
APPLICANT: Weinert, Ted A.
APPLICANT: Plon, Sharon E.
APPLICANT: Groudine, Mark T.
TITLE OF INVENTION: Cell Cycle Checkpoint Genes
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
STREET: 1420 Fifth Ave., Suite 2800
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98101-2347
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/198,446B
FILING DATE: 18-FEB-1994
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: Sheiness, Diana K.
REGISTRATION NUMBER: 35,356
REFERENCE/DOCKET NUMBER: FHCR17537
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-682-8100
TELEFAX: 206-224-0779
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 2368 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DESCRIPTION: Yeast MEC1 protein
US-08-198-446B-15

Query Match   85.7%;   Score 30;   DB 1;   Length 2368;

Best Local Similarity . 71.4%;   Pred. No. 1.9e+02;
Matches 5;   Conservative 2;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 TLLEYRM 7
           |:|:|||
Db   1647 TILDYRM 1653

RESULT 2
US-08-870-693-15
Sequence 15, Application US/0870693
Patent No. 5866338
GENERAL INFORMATION:
;   APPLICANT: Hartwell, Leland H.
;   APPLICANT: Weinert, Ted A.
;   APPLICANT: Plon, Sharon E.
;   APPLICANT: Groudine, Mark T.
;   TITLE OF INVENTION: Cell Cycle Checkpoint Genes
;   NUMBER OF SEQUENCES: 19
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Christensen O'Connor Johnson Kindness PLLC
;   STREET: 1420 Fifth Ave., Suite 2800
;   CITY: Seattle
;   STATE: WA
;   COUNTRY: USA
;   ZIP: 98101-2347
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/870,693
;   FILING DATE:
;   PRIOR APPLICATION DATA:
;   CLASSIFICATION: 536
;   APPLICATION NUMBER: US 08/198,446
;   FILING DATE: February 18, 1994
;   APPLICATION NUMBER: PCT/US93/04458
;   FILING DATE: May 12, 1993
;   APPLICATION NUMBER: US 07/884,426
;   FILING DATE: May 14, 1992
;   APPLICATION NUMBER: US 07/882,051
;   FILING DATE: May 12, 1992
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Sheiness, Diana K.
;   REGISTRATION NUMBER: 35,356
;   REFERENCE/DOCKET NUMBER: FHCR110798
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: 206-682-8100
;   TELEFAX: 206-224-0779
;   INFORMATION FOR SEQ ID NO: 15:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2368 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   DESCRIPTION: Yeast MEC1 protein
US-08-870-693-15

Qy   1 TLLEYRM 7
           |:|:|||
Db   1647 TILDYRM 1653

RESULT 3
US-08-286-819A-19
; Sequence 19, Application US/08286819A

Query Match   85.7%;   Score 30;   DB 2;   Length 2368;
Best Local Similarity   71.4%;   Pred. No. 1.9e+02;
Matches 5;   Conservative 2;   Mismatches 0;   Indels 0;   Gaps 0;

Patent No. 5871910
GENERAL INFORMATION:
;   APPLICANT: ARTHUR, MICHEL
;   APPLICANT: DUKTA-MALEN, SYLVIE
;   APPLICANT: MOLINAS, CATHERINE
;   APPLICANT: COURVALIN, PATRICE
;   TITLE OF INVENTION: POLYPEPTIDES IMPLICATED IN THE
;   TITLE OF INVENTION: EXPRESSION OF RESISTANCE TO GLYCOPEPTIDES, IN PARTICULAR
;   TITLE OF INVENTION: IN GRAM-POSITIVE BACTERIA, NUCLEOTIDE SEQUENCE CODING FOR
;   TITLE OF INVENTION: THESE POLYPEPTIDES AND USE FOR DIAGNOSIS
;   NUMBER OF SEQUENCES: 54
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
;   ADDRESSEE: P.C.
;   STREET: 1755 S. Jefferson Davis Highway, Suite 400
;   CITY: Arlington
;   STATE: Virginia
;   COUNTRY: U.S.A.
;   ZIP: 22202
;   COMPUTER READABLE FORM:
;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/286,819A
;   FILING DATE: 05-AUG-1994
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 08/174,682
;   FILING DATE: 28-DEC-1993
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: US 07/917,146
;   FILING DATE: 10-AUG-1992
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: PCT/FR/91/00855
;   FILING DATE: 29-OCT-1991
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: FR 9013579
;   FILING DATE: 31-OCT-1990
;   CLASSIFICATION: 435
;   ATTORNEY/AGENT INFORMATION:
;   NAME: Oblon, No. 587191010man F.
;   REGISTRATION NUMBER: 24,618
;   REFERENCE/DOCKET NUMBER: 660-060-0 PCT
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (703) 413-3000
;   TELEFAX: (703) 413-2220
;   TELEX: 248855 OPAT UR
;   INFORMATION FOR SEQ ID NO: 19:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 988 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
;   MOLECULE TYPE: protein
US-08-286-819A-19

Qy   1 TLLEYRM 7
           |:|:|||
Db   120 TLSEYRM 126

RESULT 4
US-08-980-357-19
; Sequence 19, Application US/08980357
; Patent No. 6013508

Query Match   82.9%;   Score 29;   DB 2;   Length 988;
Best Local Similarity   85.7%;   Pred. No. 1.3e+02;
Matches 6;   Conservative 0;   Mismatches 1;   Indels 0;   Gaps 0;

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen. Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:18:00 ; Search time 102.84 seconds
                                              (without alignments)
                                              8.148 Million cell updates/sec

Title: US-09-281-760C-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 219241

1mum DB seq length: 0
1mum DB seq length: 200000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   PIR_68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

          Pred. No. is the number of results predicted by chance to have a
          score greater than or equal to the score of the result being printed,
          and is derived by analysis of the total score distribution.

                              SUMMARIES

Result No.	Score	% Query Match	Length	Match DB	ID	Description
1	42	66.7	795	1	S73830	endopeptidase La (
2	41	65.1	351	1	SAVZVV	surface antigen pr
3	41	65.1	351	1	SAVZWR	surface antigen pr
4	41	65.1	353	1	SAVZVC	surface antigen pr
5	41	65.1	354	2	S46877	B20R protein - var
6	41	65.1	354	2	T28616	hypothetical prote
7	41	65.1	355	2	D7Z174	D9R protein - vari
8	41	65.1	583	2	I39428	alcam - human
9	41	65.1	367	2	B31847	hypothetical prote
10	39	61.9	368	2	JT0348	VSG expression sit
11	39	61.9	1014	2	S75724	hypothetical prote
12	39	61.9	129	1	I54411	MHC RT1-B A-alpha
13	38	60.3	221	1	HLMSA1	H-2 class II histo
14	38	60.3	233	1	I79358	IA-alpha polyprote
15	38	60.3	233	2	I79357	IA-alpha polyprote
16	38	60.3	254	2	S16649	class II histocomp
17	38	60.3	256	1	HLMSA2	H-2 class II histo
18	38	60.3	256	2	S02855	class II histocomp
19	38	60.3	256	2	I54447	gene MHC H2-I-A-al
20	38	60.3	258	1	HLMSAA	H-2 class II histo
21	38	60.3	310	2	A86812	sugar ABC transpor
22	38	60.3	454	2	A46532	Ig mu chain C regi
23	38	60.3	490	1	A55141	GlcNAc beta-1,4-N-
24	38	60.3	526	2	A37821	butyrophilin - bov
25	38	60.3	530	2	S00163	glucuronosyltransf
26	38	60.3	555	2	JQ1526	interleukin-1 rece
27	38	60.3	754	2	S62512	probable cysteinyl
28	38	60.3	26926	1	I38344	titin, cardiac mus
29	37	58.7	235	2	F64799	hypothetical prote
30	37	58.7	235	2	C85565	probable tRNA liga
31	37	58.7	691	2	D84889	probable receptor-
32	37	58.7	887	2	A46764	microsomal triglyc
33	37	58.7	894	2	I38047	triglyceride trans
34	37	58.7	895	2	A55413	triglyceride trans
35	37	58.7	1291	2	T09273	probable tail-host
36	37	58.7	1597	2	S65053	genome polyprotein
37	36	57.1	81	2	D81873	hypothetical prote
38	36	57.1	115	2	B46518	hypothetical prote
39	36	57.1	182	2	A25468	probable lipoprote
40	36	57.1	240	2	S01299	hypothetical prote
41	36	57.1	391	2	H85063	hypothetical prote
42	36	57.1	447	2	T34992	probable cytochrom
43	36	57.1	515	2	S54592	hypothetical prote
44	36	57.1	520	2	H84663	probable cytochrom
45	36	57.1	577	2	I50731	Ig heavy chain - n
46	36	57.1	628	2	C82039	conserved hypothet
47	36	57.1	729	2	A49120	fibroblast growth
48	36	57.1	776	2	S13470	isoamylase (EC 3.2
49	36	57.1	776	2	A37035	isoamylase (EC 3.2
50	36	57.1	837	2	T19825	hypothetical prote

                              ALIGNMENTS

RESULT   1
S73830
endopeptidase La (EC 3.4.21.53) - Mycoplasma pneumoniae (strain ATCC 29342)
N;Alternate names: ATP-dependent proteinase lon; ATP-dependent serine proteinase La;
N;Contains: adenosinetriphosphatase (EC 3.6.1.3)
C;Species: Mycoplasma pneumoniae
A;Variety: ATCC 29342
C;Date: 09-Apr-1999 #sequence_revision 09-Apr-1999 #text_change 19-Jan-2001
C;Accession: S73830
R;Himmelreich, R.; Hilbert, H.; Plagens, H.; Pirkl, E.; Li, B.C.; Herrmann, R.
Nucleic Acids Res. 24, 4420-4449, 1996
A;Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon
A;Reference number: S73327; MUID:97105885
A;Accession: S73830
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-795 <HIM>
A;Cross-references: EMBL:AE000050; GB:U00089; NID:g1674197; PIDN:AAB96152.1; PID:g167
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1996
C;Comment: This enzyme catalyzes the hydrolysis of large proteins in the presence of
C;Genetics:
A;Gene: lon; F10_orf795
A;Genetic code: SGC3
C;Superfamily: ATP-dependent serine proteinase La
C;Keywords: ATP; DNA binding; heat shock; hydrolase; nucleotide binding; P-loop; seri
F;379-386/Region: nucleotide-binding motif A (P-loop)
F;442-447/Region: nucleotide-binding motif B
F;702/Active site: Ser #status predicted

Query Match   66.7%;   Score 42; DB 1; Length 795;
Best Local Similarity 70.0%;   Pred. No. 18;
   Matches 7; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy   2 MNLTWYRESK 11
           || |:|||||
Db   322 MNLPWWQESK 331

RESULT   2
SAVZVV
surface antigen precursor - vaccinia virus
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jun-2000
C;Accession: A35522
R;Ueda, Y.; Morikawa, S.; Matsuura, Y.

RESULT 3
SAVZWR
surface antigen precursor - vaccinia virus (strain WR)
N;Alternate names: B18R protein
A;Accession: B38472
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
A;Molecule type: DNA
A;Residues: 1-351 <SM1>
A;Cross-references: GB:D01019; NID:g222698; PIDN:BAA00826.1; PID:g222699
R;Smith, G.L.; Chan, Y.S.; Howard, S.T.
J;Gen. Virol. 72, 1349-1376, 1991
A;Title: Nucleotide sequence of 42kbp of vaccinia virus strain WR from near the right in
A;Reference number: JQ1767; MUID:91250063
A;Accession: JQ1812
R;Smith, G.L.; Chan, Y.S.
J;Gen. Virol. 72, 511-518, 1991
A;Title: Two vaccinia virus proteins structurally related to the interleukin-1 receptor
A;Reference number: A38472; MUID:91170931
A;Accession: B38472
A;Molecule type: DNA
A;Residues: 1-351 <SM2>
;Cross-references: DDBJ:D11079; NID:g222717; PIDN:BAA01848.1; PID:g222763; DDBJ:D01019;
A;Reference number: A35522; MUID:90320131
A;Title: Identification and nucleotide sequence of the gene encoding a surface antigen i
Virology 177, 588-594, 1990
A;Accession: A35522
A;Molecule type: mRNA
A;Residues: 1-351 <UED>
A;Cross-references: EMBL:D90076; NID:g222769; PIDN:BAA14116.1; PID:g222770
C;Superfamily: vaccinia virus surface antigen; immunoglobulin homology
C;Keywords: glycoprotein; surface antigen
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-351/Product: surface antigen #status predicted <SAT>
F;66-131/Domain: immunoglobulin homology <IMM1>
F;165-223/Domain: immunoglobulin homology <IMM2>
F;265-335/Domain: immunoglobulin homology <IMM3>
F;73-129,172-221,272-333/Disulfide bonds: #status predicted
F;117,182,261,269/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Query Match             65.1%;  Score 41;  DB 1;  Length 351;
Best Local Similarity   55.6%;  Pred. No. 12;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3 NLTWYRESK 11
         |:|||:::|
Db 182 NITWYKDNK 190
```

RESULT 4
SAVZVC
surface antigen precursor - vaccinia virus (strain Copenhagen)
C;Species: vaccinia virus
A;Note: host Homo sapiens (man)
C;Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 21-Jan-2000
C;Accession: 142527
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.
Virology 179, 517-563, 1990
A;Title: Appendix to "The complete DNA sequence of vaccinia virus".
A;Reference number: A42501
A;Accession: 142527
A;Molecule type: DNA
A;Residues: 1-353 <GOE>
A;Cross-references: GB:M35027; NID:g335317; PIDN:AAA48218.1; PID:g335566
R;Goebel, S.J.; Johnson, G.P.; Perkus, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti,
Virology 179, 247-266, 1990
A;Title: The complete DNA sequence of vaccinia virus.
A;Reference number: A42531; MUID:91021027
A;Contents: annotation; possible protein-coding frames
A;Note: neither amino acid nor nucleotide sequence is given
C;Superfamily: vaccinia virus surface antigen; immunoglobulin homology
C;Keywords: glycoprotein; surface antigen
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-353/Product: surface antigen #status predicted <SAT>
F;267-337/Domain: immunoglobulin homology <IMM>
F;119,184,263,271/Binding site: carbohydrate (Asn) (covalent) #status predicted

```
Query Match             65.1%;  Score 41;  DB 1;  Length 351;
Best Local Similarity   55.6%;  Pred. No. 12;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3 NLTWYRESK 11
         |:|||:::|
Db 184 NITWYKDNK 192
```

RESULT 5
S46877
B20R protein - variola virus
N;Alternate names: B19R protein (COP)
C;Species: variola virus
A;Molecule type: DNA
A;Accession: S46877
A;Status: preliminary
A;Reference number: S46877
A;Accession: S46868
A;Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
submitted to the EMBL Data Library, April 1992
R;Blinov, V.M.; Gytorov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
R;Kolykhalov, A.A.; Blinov, V.M.;
C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Mar-2001
A;Accession: S46877; B36857
submitted to GenBank, November 1992
A;Reference number: A36859
A;Accession: B36857
A;Molecule type: DNA
A;Residues: 1-354 <KOL>
A;Residues: 1-354 <BLI>
A;cross-references: GB:X69198; NID:g456758; PIDN:CAA49129.1; PID:g457079
A;Experimental source: strain India-1967, ssp. major, isolate Ind3
A;Experimental source: strain India-1967, isolate Ind3
A;Cross-references: EMBL:X67117; NID:g516428; PIDN:CAA47529.1; PID:g516438
C;Superfamily: vaccinia virus surface antigen; immunoglobulin homology
F;268-338/Domain: immunoglobulin homology homology

```
Query Match             65.1%;  Score 41;  DB 2;  Length 354;
Best Local Similarity   55.6%;  Pred. No. 12;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3 NLTWYRESK 11
         |:|||:::|
Db 185 NITWYKDNK 193
```

RESULT 6

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:　　　　　April 21, 2002, 10:37:27 ; Search time 85.5 Seconds
　　　　　　　　　　　　　　　　　　　　　　　　　(without alignments)
　　　　　　　　　　　　　　　　　　　　　　　　　4.717 Million cell updates/sec

Title:　　　　　US-09-281-760C-5
Perfect score:　63
Sequence:　　　1 GMNLTWYRESK 11
Scoring table:　BLOSUM62
　　　　　　　　　Gapop 10.0 , Gapext 0.5

Database :　　　SwissProt_39:*

Searched:　　　100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:　　100059

　　inum DB seq length: 0
　　inum DB seq length: 200000000000

Post-processing: Minimum Match 0%
　　　　　　　　Maximum Match 100%
　　　　　　　　Listing first 50 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	42	66.7	795	1	LON_MYCPN	P78025 mycoplasma
2	41	65.1	351	1	VB19_VACCD	P23998 vaccinia vi
3	41	65.1	351	1	VB19_VACCV	P25213 vaccinia vi
4	41	65.1	351	1	VB19_VARV	P33795 variola vir
5	41	65.1	353	1	VB19_VACCC	P21077 vaccinia vi
6	41	65.1	583	1	C166_HUMAN	Q13740 homo sapien
7	41	65.1	583	1	C166_MOUSE	Q51490 mus musculu
8	41	65.1	221	1	HA2O_MOUSE	P04227 mus musculu
9	38	60.3	227	1	HA2U_MOUSE	P14438 mus musculu
10	38	60.3	233	1	HA2F_MOUSE	P14435 mus musculu
11	38	60.3	233	1	HA2R_MOUSE	P14436 mus musculu
11	38	60.3	233	1	HA2S_MOUSE	P14437 mus musculu
12	38	60.3	254	1	HA2J_MOUSE	P23150 mus musculu
13	38	60.3	256	1	HA2B_MOUSE	P14434 mus musculu
14	38	60.3	256	1	HA2B_RAT	P20037 rattus norv
15	38	60.3	256	1	HA2D_MOUSE	P04228 mus musculu
16	38	60.3	256	1	HA2K_MOUSE	P01910 mus musculu
17	38	60.3	490	1	BAGT_LIMST	Q09323 lymnaea sta
18	38	60.3	526	1	BUTY_BOVIN	P18892 bos taurus
19	38	60.3	530	1	UDB5_MOUSE	P17717 mus musculu
20	38	60.3	754	1	YAF6_SCHPO	Q09860 schizosacch
21	37	58.7	235	1	YBEU_ECOLI	P77427 escherichia
22	37	58.7	887	1	MTP_BOVIN	P55156 bos taurus
23	37	58.7	894	1	MTP_HUMAN	P55157 homo sapien
24	37	58.7	894	1	MTP_MOUSE	O08601 mus musculu
25	37	58.7	895	1	MTP_MESAU	P55158 mesocricetu
26	37	58.7	1597	1	RRPO_CRMV	O66220 chinese rap
27	37	58.7	182	1	CD3G_HUMAN	P09693 homo sapien
28	36	57.1	240	1	CD48_RAT	P10252 rattus norv
29	36	57.1	515	1	YM8O_YEAST	Q03264 saccharomyc
30	36	57.1	515	1	FGR1_DROME	Q07407 drosophila
31	36	57.1	730	1	ISOA_PSEAY	P10342 pseudomonas
32	36	57.1	776	1	ISOA_PSESP	P26501 pseudomonas
33	36	57.1	776	1		
34	36	57.1	837	1	YLE5_CAEEL	P46941 caenorhabdi
35	35	55.6	180	1	RL5_CHLMU	Q9pjm6 chlamydia m
36	35	55.6	180	1	RL5_CHLTR	P28531 chlamydia t
37	35	55.6	370	1	HVC1_HETFR	P23084 heterodontu
38	35	55.6	393	1	HVC3_HETFR	P23086 heterodontu
39	35	55.6	406	1	YNQ5_YEAST	P53891 saccharomyc
40	35	55.6	421	1	O67B_DROME	Q9vt20 drosophila
41	35	55.6	438	1	HVC2_HETFR	P23085 heterodontu
42	35	55.6	438	1	HVC5_HETFR	P23087 heterodontu
43	35	55.6	461	1	HVCM_HETFR	P23088 heterodontu
44	35	55.6	501	1	XYLB_LACLA	Q9cfg8 lactococcus
45	35	55.6	531	1	YP60_METTM	P14933 methanobact
46	35	55.6	543	1	APPA_BACSU	P42061 bacillus su
47	35	55.6	706	1	NUCL_HUMAN	P19338 homo sapien
48	35	55.6	712	1	NUCL_RAT	P13383 rattus norv
49	35	55.6	713	1	NUCL_MESAU	P08199 mesocricetu
50	35	55.6	795	1	LON_MYCGE	P47481 mycoplasma

ALIGNMENTS

RESULT 1

ID　 LON_MYCPN　　　　　　STANDARD;　　　PRT;　　795 AA.
AC　 P78025;
DT　 01-NOV-1997 (Rel. 35, Created)
DT　 01-NOV-1997 (Rel. 35, Last sequence update)
DT　 20-AUG-2001 (Rel. 40, Last annotation update)
DE　 LON OR MPN332 OR MP504.
OS　 Mycoplasma pneumoniae.
OC　 Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC　 Mycoplasmataceae; Mycoplasma.
OX　 NCBI_TaxID=2104;
RN　 [1]
RP　 SEQUENCE FROM N.A.
RC　 STRAIN=ATCC 29342 / M129;
RX　 MEDLINE=97105885; PubMed=8948633;
RA　 Himmelreich R., Hilbert H., Plagens H., Pirkl E., Li B.-C.,
RA　 Herrmann R.;
RT　 "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT　 pneumoniae.";
RL　 Nucleic Acids Res. 24:4420-4449(1996).
-!- FUNCTION: DEGRADES SHORT-LIVED REGULATORY PROTEINS AND ABNORMAL PROTEINS
　　　IN PRESENCE OF ATP. DEGRADES THE REGULATORY PROTEINS RCSA AND
　　　SULA. HYDROLYZES TWO ATPS FOR EACH PEPTIDE BOND CLEAVED IN THE
　　　PROTEIN SUBSTRATE (BY SIMILARITY).
-!- CATALYTIC ACTIVITY: HYDROLYSIS OF LARGE PROTEINS SUCH AS GLOBIN,
　　　CASEIN AND DENATURATED SERUM ALBUMIN, IN PRESENCE OF ATP.
-!- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
-!- SUBCELLULAR LOCATION: CYTOPLASMIC.
-!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S16; ALSO KNOWN AS THE
　　　LON FAMILY OF ATP-DEPENDENT PROTEASES.

DR　 EMBL; AE000050; AAB96152.1; -.
DR　 MEROPS; S16.004; -.
DR　 InterPro; IPR003593; AAA.
DR　 InterPro; IPR001939; AAA_subfam.
DR　 InterPro; IPR001687; ATP_GTP_A.
DR　 InterPro; IPR003111; LON.
DR　 InterPro; IPR001984; Lon_endopep.
DR　 Pfam; PF00004; AAA; 1.
DR　 Pfam; PF02190; LON; 1.
DR　 PRINTS; PR00830; ENDOLAPTASE.

```
DR   SMART; SM00382; AAA; 1.
DR   SMART; SM00464; LON; 1.
DR   PROSITE; PS01046; LON_SER; 1.
KW   Hydrolase; Serine protease; ATP-binding; Complete proteome.
FT   DOMAIN       304    308
FT   NP_BIND      379    386       ATP (POTENTIAL).
FT   ACT_SITE     702    702       BY SIMILARITY.
SQ   SEQUENCE   795 AA;  90203 MW;  7E78550B2060C891 CRC64;
```

```
Qy.    2 MNLTWYRESK 11
         ||| |||:|||
Db   322 MNLPWWQESK 331
```

```
Query Match             66.7%;  Score 42;  DB 1;  Length 795;
Best Local Similarity   70.0%;  Pred. No. 6.7;
Matches   7;  Conservative   2;  Mismatches   1;  Indels   0;  Gaps   0;
```

RESULT 2
VB19_VACCD

```
ID   VB19_VACCD     STANDARD;      PRT;   351 AA.
AC   P23998;
DT   01-MAR-1992 (Rel. 21, Created)
DT   01-MAR-1992 (Rel. 21, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   SURFACE ANTIGEN S PRECURSOR (S ANTIGEN).
GN   B19R.
OS   Vaccinia virus (strain Dairen I).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10250;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90320131; PubMed=2196742;
RA   Ueda Y., Morikawa S., Matsuura Y.;
RT   "Identification and nucleotide sequence of the gene encoding a
RT   surface antigen induced by vaccinia virus.";
RL   Virology 177:588-594(1990).
CC   -!- FUNCTION: MAY BIND INTERLEUKIN-1 AND/OR INTERLEUKIN-6 AND PREVENT
CC       THESE CYTOKINES REACHING THEIR NATURAL RECEPTORS. IN CONSEQUENCE
CC       THE INFLAMMATORY RESPONSE WOULD BE DIMINISHED AND VIRUS
CC       REPLICATION ENHANCED.
CC   -!- SUBCELLULAR LOCATION: INDUCED ON THE SURFACE OF VACCINIA
CC       VIRUS-INFECTED CELLS.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: TO THE INTERLEUKIN-1 RECEPTOR.
```

```
DR   EMBL; D90076; BAA14116.1; -.
DR   PIR; A35522; SAVZVV.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; Ig; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Antigen; Glycoprotein; Signal; Early protein.
FT   SIGNAL         1     19
FT   CHAIN         20    351       SURFACE ANTIGEN S.
FT   DOMAIN        65    137       IG-LIKE C2-TYPE DOMAIN 1.
FT   DOMAIN       164    229       IG-LIKE C2-TYPE DOMAIN 2.
FT   DOMAIN       264    341       IG-LIKE V-TYPE DOMAIN.
FT   DISULFID      73    129       BY SIMILARITY.
FT   DISULFID     172    221       BY SIMILARITY.
FT   DISULFID     272    333       BY SIMILARITY.
FT   CARBOHYD     117    117       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     182    182       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     261    261       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     269    269       N-LINKED (GLCNAC...) (POTENTIAL).
FT   CARBOHYD     321    321       N-LINKED (GLCNAC...) (POTENTIAL).
SQ   SEQUENCE   351 AA;  40701 MW;  20997CB67D39E7DB CRC64;
```

```
Qy.    3 NLTWYRESK 11
         |:|||:||:|
Db   182 NITWYKDNK 190
```

```
Query Match             65.1%;  Score 41;  DB 1;  Length 351;
Best Local Similarity   55.6%;  Pred. No. 4.4;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;
```

RESULT 3
VB19_VACCV

```
ID   VB19_VACCV     STANDARD;      PRT;   351 AA.
AC   P25213;
DT   01-MAY-1992 (Rel. 22, Created)
DT   01-MAY-1992 (Rel. 22, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   SURFACE ANTIGEN S PRECURSOR (S ANTIGEN).
GN   B19R OR B18R.
OS   Vaccinia virus (strain WR).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10254;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91259063; PubMed=2045793;
RA   Smith G.L., Chan Y.S., Howard S.T.;
RT   "Nucleotide sequence of 42 kbp of vaccinia virus strain WR from near
RT   the right inverted terminal repeat.";
RL   J. Gen. Virol. 72:1349-1376(1991).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91170931; PubMed=1826022;
RA   Smith G.L., Chan Y.S.;
RT   "Two vaccinia virus proteins structurally related to the
RT   interleukin-1 receptor and the immunoglobulin superfamily.";
RL   J. Gen. Virol. 72:511-518(1991).
CC   -!- FUNCTION: MAY BIND INTERLEUKIN-1 AND/OR INTERLEUKIN-6 AND PREVENT
CC       THESE CYTOKINES REACHING THEIR NATURAL RECEPTORS. IN CONSEQUENCE
CC       THE INFLAMMATORY RESPONSE WOULD BE DIMINISHED AND VIRUS
CC       REPLICATION ENHANCED.
CC   -!- SUBCELLULAR LOCATION: INDUCED ON THE SURFACE OF VACCINIA
CC       VIRUS-INFECTED CELLS.
CC   -!- SIMILARITY: CONTAINS 2 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAINS.
CC   -!- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC   -!- SIMILARITY: TO THE INTERLEUKIN-1 RECEPTOR.
```

```
DR   EMBL; D11079; BAA01848.1; -.
DR   EMBL; D01019; BAA00826.1; -.
DR   EMBL; X56122; CAA39587.1; -.
DR   EMBL; A14979; CAA01478.1; -.
DR   PIR; B38477; SAVZWR.
DR   PIR; JQ1812; JQ1812.
DR   InterPro; IPR003006; Ig_MHC.
DR   InterPro; IPR003600; Ig_like.
DR   Pfam; PF00047; Ig; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Immunoglobulin domain; Antigen; Glycoprotein; Signal; Early protein.
FT   SIGNAL         1     19       POTENTIAL.
```

```
OM protein - protein search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             April 21, 2002, 10:35:54 ; Search time 191.29 Seconds
                                              (without alignments)
                                              8.411 Million cell updates/sec

Title:              US-09-281-760C-5
Perfect score:      63
Sequence:           1 GMNLTWYRESK 11
Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters:   473505

Post-processing:    Minimum Match 0%
    imum DB seq length: 0                    Maximum Match 100%
    imum DB seq length: 200000000            Listing first 50 summaries

Database :          SPTREMBL_17:*
                    1: sp_archea:*
                    2: sp_bacteria:*
                    3: sp_fungi:*
                    4: sp_human:*
                    5: sp_invertebrate:*
                    6: sp_mammal:*
                    7: sp_mhc:*
                    8: sp_organelle:*
                    9: sp_phage:*
                   10: sp_plant:*
                   11: sp_rodent:*
                   12: sp_virus:*
                   13: sp_vertebrate:*
                   14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                            SUMMARIES
ult                 %    Query
NO.   Score   Match Length DB    ID         Description
----- -----   ----------------   ---------- -----------
 1    42      66.7  380   2       Q9S405     proteus mir
 2    41      65.1  113   6       Q9BDE7     sus scrofa
 3    41      65.1  265   12      Q9DUN2     vaccinia vi
 4    41      65.1  315   12      Q9IFU7     vaccinia vi
 5    41      65.1  351   12      O72754     cowpox viru
 6    41      65.1  353   12      Q9JF34     vaccinia vi
 7    41      65.1  354   12      O89915     variola vir
 8    41      65.1  355   12      O89923     variola vir
 9    41      65.1  521   6       O46634     canis famil
10    41      65.1  521   6       O46651     oryctolagus
11    41      65.1  583   6       Q9BH13     bos taurus
12    41      65.1  583   11      O35112     rattus norv
13    40      63.5  541   5       Q9NA52     caenorhabdi
14    39      61.9  191   5       O76423     trypanosoma
15    39      61.9  244   5       O76422     trypanosoma
16    39      61.9  275   11      Q9JK39     mus musculu
17    39      61.9  311   2       Q9F446     streptococc
18    39      61.9  367   5       O26723     trypanosoma
19    39      61.9  368   5       O26748     trypanosoma
20    39      61.9  1014  2       O55374     synechocyst
21    39      61.9  2196  3       Q9HDV6     schizosacch
22    38      60.3  63    7       Q9TPJ8     anguilla sp
23    38      60.3  63    7       Q9TPG1     salarias sp
24    38      60.3  129   7       O31259     rattus norv
25    38      60.3  129   7       O31264     rattus norv
26    38      60.3  227   7       O31247     peromyscus
27    38      60.3  240   11      Q9DCS4     mus musculu
28    38      60.3  254   7       Q9TQ72     mus musculu
29    38      60.3  254   7       Q9TQ71     mus musculu
30    38      60.3  286   6       O46535     bos taurus
31    38      60.3  310   2       Q9CFH9     lactococcus
32    38      60.3  312   12      P94681     comamonas t
33    38      60.3  467   2       O69452     mycoplasma
34    38      60.3  4650  4       O22138     arabidopsis
35    38      60.3  26926 4       Q10466     homo sapien
36    37      58.7  252   2       P94681     comamonas t
37    37      58.7  312   12      O89122     crucifer to
38    37      58.7  467   2       O69452     mycoplasma
39    37      58.7  854   12      O22138     arabidopsis
40    37      58.7  691   10      O69452     mycoplasma
41    37      58.7  1291  9       O64295     streptococc
42    36      57.1  81    2       O64295     streptococc
43    36      57.1  151   5       Q9NBB7     drosophila
44    36      57.1  181   5       Q9BS46     homo sapien
45    36      57.1  324   1       Q9HIW9     thermoplasm
46    36      57.1  391   10      Q9S9T0     arabidopsis
47    36      57.1  447   2       O86673     streptomyce
48    36      57.1  491   12      O39819     equine rota
49    36      57.1  491   12      O39820     equine rota
50    36      57.1  491   12      O39821     equine rota

                            ALIGNMENTS

RESULT 1
Q9S405

ID    Q9S405          PRELIMINARY;      PRT;    380 AA.
AC    Q9S405;
DT    01-MAY-2000 (TrEMBLrel. 13, Created)
DT    01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT    01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE    TERA.
GN    TERA.
OS    Proteus mirabilis.
OC    Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC    Proteus.
OX    NCBI_TaxID=584;
RN    [1]
RP    SEQUENCE FROM N.A.
RC    STRAIN=S2;
RA    Toptchieva A., Bryden L.J., Taylor D.E., Hoffman P.S.;
RT    "Characterization of the chromosomal tellurite resistance gene cluster
RT    in Proteus mirabilis.";
RL    Submitted (JUL-1999) to the EMBL/GenBank/DDBJ databases.
DR    EMBL; AF168355; AAD47285.1; -.
DR    InterPro; IPR003325; TerD.
DR    Pfam; PF02342; TerD; 2.
SQ    SEQUENCE   380 AA;   41562 MW;  941B4CBC7D33B2DE CRC64;

Query Match             66.7%;  Score 42;  DB 2;  Length 380;
Best Local Similarity   70.0%;
Matches    7;  Conservative  2;  Mismatches  1;  Indels  0;  Gaps  0;

Qy          2 MNLTWYRESK 11
               |:|| :|:||||
Db        204 INLDWHRESK 213

RESULT 2
Q9BDE7
```

```
ID   Q9BDE7         PRELIMINARY;   PRT;   113 AA.
AC   Q9BDE7;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   ACTIVATED LEUKOCYTE CELL ADHESION MOLECULE (FRAGMENT).
GN   ALCAM.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=BONE MARROW;
RA   Ringe J., Kaps C., Schmitt B., Buescher K., Bartel J., Smolian H.,
RA   Schultz O., Burmester G.R., Haupl T., Sittinger M.;
RT   "Porcine Mesenchymal Stem Cells: Induction of distinct mesenchymal
RT   cell lineages.";
RL   Submitted (MAR-2001) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ311681; CAC35029.1; -.
FT   NON_TER       1      1
FT   NON_TER     113    113
SQ   SEQUENCE   113 AA;  12786 MW;  E47011249F443F CRC64;
Qy   3  NLTWYRNGK 113
        |:|||||  |
Db   105  NITWYRNGK 113
```

```
RESULT  3
ID   Q9DUN2         PRELIMINARY;   PRT;   265 AA.
AC   Q9DUN2;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   INTERFERON-ALPHA/BETA RECEPTOR.
GN   B18R.
OS   Vaccinia virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10245;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WYETH;
RX   PubMed=11070021;
RA   Alcami A., Symons J.A., Smith G.L.;
RT   "The vaccinia virus soluble alpha/beta interferon (IFN) receptor binds
RT   to the cell surface and protects cells from the antiviral effects of
RT   IFN.";
RL   J. Virol. 74:11230-11239(2000).
-!-  SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
     DOMAIN.
DR   EMBL; AJ269556; CAC16604.1; -.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 1.
DR   SMART; SM00410; IG_like; 1.
KW   Receptor.
SQ   SEQUENCE   265 AA;  31038 MW;  BA43AEEB1DAA1A49 CRC64;

Query Match             65.1%;  Score 41;  DB 12;  Length 265;
Best Local Similarity   55.6%;  Pred. No. 12;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3  NLTWYRNGK 11
        |:|||||::|
Db   184  NITWYKDNK 192
```

```
RESULT  4
ID   Q9IFU7         PRELIMINARY;   PRT;   315 AA.
AC   Q9IFU7;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   SOLUBLE IFN RECEPTOR (FRAGMENT).
OS   Bean 58058 virus.
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=67082;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Marques J.T., DaFonseca F.G., Kroon E.G.;
RT   "The genome of Bean 58058 virus contains a gene homologous to Vaccinia
RT   virus B18R gene.";
RL   Submitted (APR-2000) to the EMBL/GenBank/DDBJ databases.
-!-  SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
     DOMAIN.
DR   EMBL; AF261890; AAF74754.1; -.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
DR   Pfam; PF00047; ig; 2.
DR   SMART; SM00410; IG_like; 1.
KW   Receptor.
FT   NON_TER       1      1
FT   NON_TER     315    315
SQ   SEQUENCE   315 AA;  36544 MW;  AAAC6CBDA69095AA CRC64;

Query Match             65.1%;  Score 41;  DB 12;  Length 315;
Best Local Similarity   55.6%;  Pred. No. 14;
Matches   5;  Conservative   4;  Mismatches   0;  Indels   0;  Gaps   0;

Qy   3  NLTWYRNGK 11
        |:|||||::|
Db   157  NITWYKDNK 165
```

```
RESULT  5
ID   O72754         PRELIMINARY;   PRT;   351 AA.
AC   O72754;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE   B17R PROTEIN.
GN   B17R.
OS   Cowpox virus (CPV).
OC   Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;
OC   Orthopoxvirus.
OX   NCBI_TaxID=10243;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GRI-90;
RX   MEDLINE=98229462; PubMed=9560042;
RA   Shchelkunov S.N., Safronov P.F., Totmenin A.V., Petrov N.A.,
RA   Ryazankina O.I., Gutorov V.V., Kotwal G.J.;
RT   "The genomic sequence analysis of the left and right species-specific
RT   terminal region of a cowpox virus strain reveals unique sequences and
RT   a cluster of intact ORFs for immunomodulatory and host range
RT   proteins.";
RL   Virology 243:432-460(1998).
-!-  SIMILARITY: TO IMMUNOGLOBULIN AND MAJOR HISTOCOMPATIBILITY COMPLEX
     DOMAIN.
DR   EMBL; Y15035; CAA75293.1; -.
DR   InterPro; IPR003599; Ig.
DR   InterPro; IPR003600; Ig_like.
DR   InterPro; IPR003006; Ig_MHC.
```

```
                                    GenCore version 4.5

                                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            April 21, 2002, 10:38:47 ; Search time 23.48 Seconds
                                                (without alignments)
                                                4.792 Million cell updates/sec

Title:             US-09-281-760C-2
Perfect score:     18
Sequence:          1 YRXXL 5

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters:       16280

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 50 summaries

Database :         Issued_Patents_AA.*
                   1:  /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                   2:  /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                   3:  /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                   4:  /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                   5:  /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                   6:  /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

Minimum DB seq length: 0
Maximum DB seq length: 5

SUMMARIES

```
Result          Query                                      %
No.   Score   Match Length   DB   ID                     DB   Description
-------------------------------------------------------------------------------
  1    14      77.8       5   2   US-08-968-676-25        Sequence 25, Appl
  2    12      66.7       4   2   US-09-214-614-2         Sequence 2, Appli
  3    12      66.7       4   2   US-08-968-676-27        Sequence 27, Appl
  4    12      66.7       4   1   US-08-240-712-4         Sequence 4, Appli
  5    12      66.7       4   1   US-08-428-488-59        Sequence 59, Appl
  6    12      66.7       4   1   US-08-428-488-60        Sequence 60, Appl
  7    12      66.7       4   1   US-08-428-488-61        Sequence 61, Appl
  8    12      66.7       4   1   US-08-428-488-62        Sequence 62, Appl
  9    12      66.7       4   1   US-08-428-488-63        Sequence 63, Appl
 10    12      66.7       4   1   US-08-428-488-64        Sequence 64, Appl
 11    12      66.7       4   1   US-08-428-488-65        Sequence 65, Appl
 12    12      66.7       4   1   US-08-428-488-66        Sequence 66, Appl
 12    12      66.7       4   1   US-08-428-488-67        Sequence 67, Appl
 13    12      66.7       4   1   US-08-428-488-68        Sequence 68, Appl
 14    12      66.7       4   1   US-08-428-488-69        Sequence 69, Appl
 15    12      66.7       4   1   US-08-428-488-70        Sequence 70, Appl
 16    12      66.7       4   1   US-08-428-488-71        Sequence 71, Appl
 17    12      66.7       4   1   US-08-428-488-72        Sequence 72, Appl
 18    12      66.7       4   1   US-08-428-488-73        Sequence 73, Appl
 19    12      66.7       4   1   US-08-428-488-74        Sequence 74, Appl
 20    12      66.7       4   1   US-08-255-272-22        Sequence 22, Appl
 21    12      66.7       4   1   US-08-335-198-25        Sequence 25, Appl
 22    12      66.7       4   1   US-08-429-732-20        Sequence 20, Appl
 23    12      66.7       4   1   US-08-203-806B-15       Sequence 15, Appl
 24    12      66.7       4   1   US-08-443-890-4         Sequence 4, Appl
 25    12      66.7       4   1   US-08-706-178-2         Sequence 2, Appli
 26    12      66.7       4   1   US-08-441-871-1         Sequence 1, Appli
 27    12      66.7       4   2   US-08-441-871-63        Sequence 63, Appl
 28    12      66.7       4   2   US-08-441-871-65        Sequence 65, Appl
 29    12      66.7       4   2   US-08-441-871-66        Sequence 66, Appl
 30    12      66.7       4   2   US-08-441-871-68        Sequence 68, Appl
 31    12      66.7       4   2   US-08-441-871-70        Sequence 70, Appli
 32    12      66.7       4   2   US-08-473-022-2         Sequence 2, Appli
 33    12      66.7       4   2   US-08-705-660-27        Sequence 27, Appl
 34    12      66.7       4   2   US-08-705-660-28        Sequence 28, Appl
 35    12      66.7       4   2   US-08-623-833B-61       Sequence 61, Appl
 36    12      66.7       4   2   US-08-685-589A-6        Sequence 6, Appli
 37    12      66.7       4   2   US-08-685-589A-62       Sequence 62, Appl
 38    12      66.7       4   2   US-08-968-676-26        Sequence 26, Appl
 39    12      66.7       4   2   US-08-752-844-12        Sequence 12, Appl
 40    12      66.7       4   2   US-08-747-137-78        Sequence 78, Appl
 41    12      66.7       4   2   US-08-551-196-12        Sequence 12, Appl
 42    12      66.7       4   2   US-08-989-045-27        Sequence 27, Appl
 43    12      66.7       4   3   US-08-989-045-28        Sequence 28, Appl
 44    12      66.7       4   3   US-09-047-986B-20       Sequence 20, Appl
 45    12      66.7       4   3   US-09-082-593-11        Sequence 11, Appl
 46    12      66.7       4   4   US-09-058-562-4         Sequence 4, Appli
 47    12      66.7       4   4   PCT-US92-09752-4        Sequence 4, Appli
 48    12      66.7       4   5   PCT-US93-01669-16       Sequence 16, Appl
 49    12      66.7       4   5   PCT-US96-08798-2        Sequence 2, Appli
 50    12      66.7       4                               
```

ALIGNMENTS

```
RESULT       1
US-08-968-676-25

Sequence 25, Application US/08968676

Patent No. 5919639

GENERAL INFORMATION:

   APPLICANT:  Humphreys, Robert E
   APPLICANT:  Adams, Sharlene
   APPLICANT:  Xu, Minzhen
   TITLE OF INVENTION:  IMMUNOTHERAPY BY MODULATION OF ANTIGEN
   NUMBER OF SEQUENCES:  165
   CORRESPONDENCE ADDRESS:
       ADDRESSEE:  Kevin M. Farrell, P.C.
       STREET:  P.O. Box 999
       CITY:  York Harbor
       STATE:  ME
       COUNTRY:  USA
       ZIP:  03911
   COMPUTER READABLE FORM:
       MEDIUM TYPE:  Floppy disk
       COMPUTER:  IBM PC compatible
       OPERATING SYSTEM:  PC-DOS/MS-DOS
       SOFTWARE:  Patentin Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
       APPLICATION NUMBER:  US/08/968,676
       FILING DATE:
       CLASSIFICATION:  424
   ATTORNEY/AGENT INFORMATION:
       NAME:  Farrell, Kevin M
       REGISTRATION NUMBER:  35,505
       REFERENCE/DOCKET NUMBER:  REH-9601
   TELECOMMUNICATION INFORMATION:
       TELEPHONE:  (207) 363-0558
       TELEFAX:  (207) 363-0528
   INFORMATION FOR SEQ ID NO:  25:
       SEQUENCE CHARACTERISTICS:
           LENGTH:  5 amino acids
           TYPE:  amino acid
           STRANDEDNESS:  single
           TOPOLOGY:  linear
       MOLECULE TYPE:  peptide

US-08-968-676-25

Query Match            77.8%;   Score 14;  DB 2;  Length 5;
Best Local Similarity  60.0%;   Pred. No. 1.6e+05;
```

Qy　1 YRXXL 5
　　　 | (| |
Db　1 YRMKL 5

Matches　3;　Conservative　0;　Mismatches　2;　Indels　0;　Gaps　0;

RESULT 2
US-09-214-614-2
Sequence 2, Application US/09214614
Patent No. 6225100
GENERAL INFORMATION:
　APPLICANT: Grund, Alan D.
　APPLICANT: Maurina-Brunker, Julie
　TITLE OF INVENTION: NOVEL ARYLSULFOTRANSFERASE
　NUMBER OF SEQUENCES: 10
　CORRESPONDENCE ADDRESS:
　　ADDRESSEE: Sheridan Ross P.C.
　　STREET: 1700 Lincoln Street, Suite 3500
　　CITY: Denver
　　STATE: Colorado
　　COUNTRY: U.S.A
　　ZIP: 80203
　COMPUTER READABLE FORM:
　　MEDIUM TYPE: Floppy disk
　　COMPUTER: IBM PC compatible
　　OPERATING SYSTEM: PC-DOS/MS-DOS
　　SOFTWARE: PatentIn Release #1.0, Version #1.25
　CURRENT APPLICATION DATA:
　　APPLICATION NUMBER: US/09/214,614
　　FILING DATE:
　　CLASSIFICATION: 435
　ATTORNEY/AGENT INFORMATION:
　　NAME: Connell, Gary J.
　　REGISTRATION NUMBER: 32,020
　　REFERENCE/DOCKET NUMBER: 3161-15-PCT
　TELECOMMUNICATION INFORMATION:
　　TELEPHONE: 303/863-9700
　　TELEFAX: 303/863-0223
　INFORMATION FOR SEQ ID NO: 2:
　　SEQUENCE CHARACTERISTICS:
　　　LENGTH: 2 amino acids
　　　TYPE: amino acid
　　　TOPOLOGY: linear
　　MOLECULE TYPE: peptide
US-09-214-614-2

Qy　1 YR 2
　　　 | |
Db　1 YR 2

Query Match　66.7%;　Score 12;　DB 4;　Length 2;
Best Local Similarity　100.0%;　Pred. No. 1.6e+05;
Matches　2;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

RESULT 3
US-08-968-676-27
Sequence 27, Application US/08968676
Patent No. 5919639
GENERAL INFORMATION:
　APPLICANT: Humphreys, Robert E
　APPLICANT: Adams, Sharlene
　APPLICANT: Xu, Minzhen
　TITLE OF INVENTION: IMMUNOTHERAPY BY MODULATION OF ANTIGEN
　NUMBER OF SEQUENCES: 165
　CORRESPONDENCE ADDRESS:
　　ADDRESSEE: Kevin M Farrell, P.C.
　　STREET: P.O. Box 999
　　CITY: York Harbor
　　STATE: ME
　　COUNTRY: USA
　　ZIP: 03911
　COMPUTER READABLE FORM:
　　MEDIUM TYPE: Floppy disk
　　COMPUTER: IBM PC compatible
　　OPERATING SYSTEM: PC-DOS/MS-DOS
　　SOFTWARE: PatentIn Release #1.0, Version #1.25
　CURRENT APPLICATION DATA:
　　APPLICATION NUMBER: US/08/968,676
　　FILING DATE:
　　CLASSIFICATION: 424
　ATTORNEY/AGENT INFORMATION:
　　NAME: Farrell, Kevin M
　　REGISTRATION NUMBER: 35,505
　　REFERENCE/DOCKET NUMBER: REH-9601
　TELECOMMUNICATION INFORMATION:
　　TELEPHONE: (207) 363-0528
　　TELEFAX: (207) 363-0558
　INFORMATION FOR SEQ ID NO: 27:
　　SEQUENCE CHARACTERISTICS:
　　　LENGTH: 3 amino acids
　　　TYPE: amino acid
　　　STRANDEDNESS: single
　　　TOPOLOGY: linear
　　MOLECULE TYPE: peptide
US-08-968-676-27

Qy　1 YR 2
　　　 | |
Db　1 YR 2

Query Match　66.7%;　Score 12;　DB 2;　Length 3;
Best Local Similarity　100.0%;　Pred. No. 1.6e+05;
Matches　2;　Conservative　0;　Mismatches　0;　Indels　0;　Gaps　0;

RESULT 4
US-08-240-712-4
Sequence 4, Application US/08240712
Patent No. 5559907
GENERAL INFORMATION:
　APPLICANT: ANDERSON, DAVID C.
　APPLICANT: MATHEWS, ANTONY JAMES
　APPLICANT: STETLER, GARY L.
　TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
　TITLE OF INVENTION: HEMOGLOBINS
　NUMBER OF SEQUENCES: 35
　CORRESPONDENCE ADDRESS:
　　ADDRESSEE: Browdy and Neimark
　　STREET: 419 Seventh Street, N.W., Suite 300
　　CITY: Washington
　　STATE: D.C.
　　COUNTRY: USA
　　ZIP: 20004
　COMPUTER READABLE FORM:
　　MEDIUM TYPE: Floppy disk
　　COMPUTER: IBM PC compatible
　　OPERATING SYSTEM: PC-DOS/MS-DOS
　　SOFTWARE: PatentIn Release #1.0, Version #1.25
　CURRENT APPLICATION DATA:
　　APPLICATION NUMBER: US/08/240,712
　　FILING DATE: 09-MAY-1994
　　CLASSIFICATION: 530
　PRIOR APPLICATION DATA:
　　APPLICATION NUMBER: PCT/US92/09752
　　FILING DATE: 13-MAY-1993
　ATTORNEY/AGENT INFORMATION:
　　NAME: COOPER, IVER P
　　REGISTRATION NUMBER: 28,005
　　REFERENCE/DOCKET NUMBER: ANDERSON-6
　TELECOMMUNICATION INFORMATION:
　　TELEPHONE: 202-628-5197
　　TELEFAX: 202-737-3528

us-09-281-760c-2.closed.rag

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:38:11 ; Search time 40.45 Seconds
(without alignments)
9.156 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues
Total number of hits satisfying chosen parameters: 20141

Post-processing: Minimum Match 08%
Maximum Match 100%
Listing first 50 summaries

imum DB seq length: 0
imum DB seq length: 5

Database : A_Geneseq_1101:*
1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Length	DB	ID	Description
1	14	77.8	5	2	AAP10205	Leu-enkephalin.
2	14	77.8	5	11	AAR08321	Peptide with opioi
3	14	77.8	5	15	AAR55499	Opioid oligopeptid
4	14	77.8	5	15	AAR55508	Opioid oligopeptid
5	14	77.8	5	15	AAR55490	Opioid oligopeptid
6	14	77.8	5	19	AAW47949	Opioid oligopeptid
7	14	77.8	5	21	AAY69582	AE101 analogue eff
8	13	72.2	5	22	AAB81468	Generic immunogeni
9	13	72.2	5	22	AAB91738	Synthetic peptide,
10	12	66.7	5	18	AAW33241	Opioid peptide SEQ
11	12	66.7	3	19	AAW56224	Anti-inflammatory
12	12	66.7	3	22	AAB91597	Opioid peptide SEQ
13	12	66.7	4	5	AAP40467	Enkephalin analogu
14	12	66.7	4	6	AAP50336	Analgesic enkephal
15	12	66.7	4	6	AAP50690	Analgesic enkephal
16	12	66.7	4	7	AAP61491	Analgesic tetrapep
17	12	66.7	4	7	AAP61741	Analgesic peptide.
18	12	66.7	4	8	AAP71341	Analgesic peptide.
19	12	66.7	4	8	AAP71580	Growth hormone rel
20	12	66.7	4	8	AAP71582	Growth hormone rel
21	12	66.7	4	8	AAP71700	Growth hormone rel
22	12	66.7	4	13	AAR23084	Analgesic peptide
23	12	66.7	4	13	AAR24410	Fluorinated peptid
24	12	66.7	4	14	AAR41643	CPase B-like enzym
25	12	66.7	4	14	AAR33660	Internalisation si
26	12	66.7	4	14	AAR33664	Growth hormone rel
27	12	66.7	4	14	AAR38136	Protease-non-labil
28	12	66.7	4	15	AAW51593	Peptide analogue #
29	12	66.7	4	15	AAW51589	Peptide analogue #
30	12	66.7	4	15	AAW51595	Peptide analogue #
31	12	66.7	4	15	AAR55497	Opioid oligopeptid
32	12	66.7	4	15	AAR55498	Opioid oligopeptid
33	12	66.7	4	15	AAR55502	Opioid oligopeptid
34	12	66.7	4	15	AAR55503	Opioid oligopeptid
35	12	66.7	4	15	AAR55506	Opioid oligopeptid
36	12	66.7	4	15	AAR55507	Opioid oligopeptid
37	12	66.7	4	15	AAR55481	Opioid oligopeptid
38	12	66.7	4	15	AAR55484	Opioid oligopeptid
39	12	66.7	4	15	AAR55485	Opioid oligopeptid
40	12	66.7	4	15	AAR55488	Opioid oligopeptid
41	12	66.7	4	15	AAR55489	Opioid oligopeptid
42	12	66.7	4	15	AAR55493	Opioid oligopeptid
43	12	66.7	4	15	AAR55494	Opioid oligopeptid
44	12	66.7	4	15	AAR55465	Opioid oligopeptid
45	12	66.7	4	15	AAR55467	Opioid oligopeptid
46	12	66.7	4	15	AAR55469	Opioid oligopeptid
47	12	66.7	4	15	AAR55471	Opioid oligopeptid
48	12	66.7	4	15	AAR55473	Opioid oligopeptid
49	12	66.7	4	15	AAR55475	Opioid oligopeptid
50	12	66.7	4	15	AAR55477	Opioid oligopeptid

ALIGNMENTS

RESULT 1
AAP10205

ID   AAP10205 standard; Protein; 5 AA.
XX
AC   AAP10205;
XX
DT   16-DEC-1992 (first entry)
XX
DE   Leu-enkephalin.
XX
KW   Analgesic; enkephalin.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   Misc-difference 2
FT   /note= "D- or L-form residue"
XX
PN   JP56045450-A.
XX
PD   25-APR-1981.
XX
PF   21-SEP-1979; 79JP-0121654.
XX
PR   21-SEP-1979; 79JP-0121654.
XX
PA   (DAUC) DAIICHI SEIYAKU KK.
XX
DR   WPI; 1981-43030D/24 (43030D).

XX
PT   Analgesic penta:peptide - comprising tyrosine, arginine, glycine,
PT   phenylalanine with methionine or leucine hydroxyl or amine
XX
PS   Disclosure; Page 3; 7pp; Japanese.
XX
CC   The peptide is a leucine-enkephalin analogue, opt. having the
CC   C-terminal amidated. It is an analgesic having stronger activity
CC   compared with the enkephalins.
CC   It is prepd. by conventional chemical synthesis, either in liq.
CC   phase or solid phase.
XX
SQ   Sequence   5 AA;

Query Match   77.8%;   Score 14;   DB 2;   Length 5;
Best Local Similarity   60.0%;   Pred. No. 4.3e+05;
Matches   3;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 YRXXL 5
            | ||
Db   1 yrgfl 5

RESULT   2

ID   AAR08321   AAR08321 standard; protein; 5 AA.
XX
AC   AAR08321;
XX
DT   05-FEB-1991   (first entry)
XX
DE   Peptide with opioid activity.
XX
KW   Analgesic.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT
FT   Modified-site   2
FT   /label= D-Arg
FT   Misc-difference   4
FT   /label= Phe or Phe(NO2) or MePhe
FT   Misc-difference   5
FT   /label= -NH-CH-CH2-CH(CH3)2-X or -Leu-NH-CH2-X
FT   or -NH-CH2-X
FT   /note=
FT   "X = -N-(2-(3-nitro-2-pyridyldithio)ethyl)
FT   amide"
XX
PN   JP02233694-A.
XX
PD   17-SEP-1990.
XX
PF   07-MAR-1989;   89JP-0052901.
XX
PR   07-MAR-1989;   89JP-0052901.
XX
PA   (VITA-) VITAMINE KENKYUSHO.
XX
DR   WPI; 1990-325185/43.
XX
PT   New peptide and its pharmaceutically acceptable salts - has
PT   opioid action, useful for analgesics
XX
PS   Claim 1; Page 861; 5pp; Japanese.
XX
CC   Peptide has opioid activity, useful in analgesics.
XX
SQ   Sequence   5 AA;

Query Match   77.8%;   Score 14;   DB 11;   Length 5;
Best Local Similarity   60.0%;   Pred. No. 4.3e+05;
.

Matches   3;   Conservative   0;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 YRXXL 5
            | ||
Db   1 yrgfl 5

RESULT   3

ID   AAR55499   AAR55499 standard; peptide; 5 AA.
XX
AC   AAR55499;
XX
DT   19-DEC-1994   (first entry)
XX
DE   Opioid oligopeptide.
XX
KW   opioid receptor; analgesic; antitussive; antidiarrhoeal; antitumour;
KW   pain killer.
XX
OS   Synthetic.
XX
FH   Key   Location/Qualifiers
FT   Modified-site   2
FT   /label= OTHER
FT   /note= "D-arginal"
FT   Modified-site   3
FT   /label= OTHER
FT   /note= "3,4-dihydroxyphenylalanine (DOPA)"
FT   Modified-site   4..5
FT   /note= "Phe-psi(CH2NH)-Leu-NH2"
FT   Modified-site   5
FT   /note= "amidated"
XX
PN   WO9411018-A.
XX
PD   26-MAY-1994.
XX
PF   04-NOV-1993;   93WO-US10711.
XX
PR   12-NOV-1992;   92US-0974943.
XX
PA   (BIOM-) BIOMEASURE INC.
XX
PI   Kim SH, Moreau J, Taylor JE;
XX
DR   WPI; 1994-183155/22.
XX
PT   New opioid oligo:peptide cpds. - having analgesic, antitussive,
PT   antidiarrhoeal and antitumour activity
XX
PS   Claim 16; Page 39; 45pp; English.
XX
CC   This peptide is an example of a highly generic formula representing
CC   opioid oligopeptides which selectively bind to receptors on cells.
CC   The peptides have analgesic, antitussive and antidiarrhoeal activity.
CC   They can also be used for treating various cancers such as lung or
CC   breast cancer, melanoma or neuroblastoma, and as tools for detecting
CC   specific opioid receptors.
XX
SQ   Sequence   5 AA;

Query Match   77.8%;   Score 14;   DB 15;   Length 5;
Best Local Similarity   80.0%;   Pred. No. 4.3e+05;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 YRXXL 5
            | || |
Db   1 yrxfl 5

RESULT   4

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:44:25 ; Search time 37.83 Seconds
                     (without alignments)
                     19.333 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5
Scoring table: BLOSUM62
                     Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 8

nimum DB seq length: 0
ximum DB seq length: 5

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   SPTREMBL_17:*
         1:    sp_archea:*
         2:    sp_bacteria:*
         3:    sp_fungi:*
         4:    sp_human:*
         5:    sp_invertebrate:*
         6:    sp_mammal:*
         7:    sp_mhc:*
         8:    sp_organelle:*
         9:    sp_phage:*
         10:   sp_plant:*
         11:   sp_rodent:*
         12:   sp_virus:*
         13:   sp_vertebrate:*
         14:   sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	% Query Match	Length	DB	ID	Description
1	22.2	4	11		Q08433	rattus norv
2	16.7	5	13		P82070	litoria rub
3	16.7	5	13		P82071	litoria rub
4	16.7	5	13		P82072	litoria rub
5	16.7	5	13		P82073	litoria rub
6	16.7	5	13		P82099	litoria rub
7	16.7	5	13		P82100	litoria rub
8	11.1	5	10		Q99007	hordeum vul

ALIGNMENTS

RESULT 1
Q08433   PRELIMINARY;   PRT;   4 AA.

ID   Q08433   PRELIMINARY;   PRT;   4 AA.
AC   Q08433;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   UDP-GLUCURONOSYLTRANSFERASE, MICROSOMAL (EC 2.4.1.17) (UDPGT)
DE   (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=91282758; PubMed=1840486;
RA   Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT   "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT   hyperbilirubinemic Gunn rat.";
RL   Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC   -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC   SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC   ENDOGENOUS COMPOUNDS.
CC   -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC   BETA-D-GLUCURONOSIDE.
CC   -!- SUBCELLULAR LOCATION: MICROSOME.
DR   EMBL; S38635; AAB19259.1; -.
KW   Transferase; Glycosyltransferase; Microsome; Multigene family.
FT   NON_TER   1   1
FT   NON_TER   4   4
SQ   SEQUENCE   4 AA;   473 MW; 633732C420000000 CRC64;

Query Match   22.28%; Score 4; DB 11; Length 4;
Best Local Similarity   100.0%; Pred. No. 4.7e+05;
Matches   1; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

Qy   5 L 5
              |
Db   3 L 3

RESULT 2
P82070   PRELIMINARY;   PRT;   5 AA.

ID   P82070   PRELIMINARY;   PRT;   5 AA.
AC   P82070;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 1.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella', the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC   ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   5 AA;   598 MW; 6DD9C9CABA2A00000 CRC64;

Query Match   16.78%; Score 3; DB 13; Length 5;
Best Local Similarity   0.0%; Pred. No. 4.7e+05;
Matches   0; Conservative   1; Mismatches   0; Indels   0; Gaps   0;

Qy   1 Y 1
              :

```
RESULT 3

Db         3 F 3

ID   P82071                 PRELIMINARY;      PRT;      5 AA.
AC   P82071;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 2.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella'. the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW   Amphibian skin; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;   626 MW;  6D09C9CB10300000 CRC64;

Qy          1 Y 1
            : .
Db          3 F 3

Query Match             16.7%;  Score 3;  DB 13;   Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
Matches    0; Conservative    1; Mismatches    0; Indels    0; Gaps    0;
```

```
RESULT 4

ID   P82072                 PRELIMINARY;      PRT;      5 AA.
AC   P82072;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 3.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella'. the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian skin; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;   656 MW;  71A9C9CB10300000 CRC64;

Query Match             16.7%;  Score 3;  DB 13;   Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
Matches    0; Conservative    1; Mismatches    0; Indels    0; Gaps    0;

Qy          1 Y 1
            : .
Db          3 F 3
```

```
RESULT 5

Db         3 F 3

ID   P82073                 PRELIMINARY;      PRT;      5 AA.
AC   P82073;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litoria electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:0-0(1999).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW   Amphibian skin.
SQ   SEQUENCE   5 AA;   570 MW;  71A9C9C862A00000 CRC64;

Qy          1 Y 1
            : .
Db          3 F 3

Query Match             16.7%;  Score 3;  DB 13;   Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
Matches    0; Conservative    1; Mismatches    0; Indels    0; Gaps    0;
```

```
RESULT 6

Db         3 F 3

ID   P82099                 PRELIMINARY;      PRT;      5 AA.
AC   P82099;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ELECTRIN 3.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litoria electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:0-0(1999).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW   Amphibian skin; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;   630 MW;  668761F2C9A00000 CRC64;

Qy          1 Y 1
            : .
Db          3 F 3

Query Match             16.7%;  Score 3;  DB 13;   Length 5;
```

```
GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           April 21, 2002, 10:44:57 ; Search time 17.3 Seconds
                                             (without alignments)
                                             10.597 Million cell updates/sec

Title:            US-09-281-760c-2
Perfect score:    18
Sequence:         1 YRXXL 5

Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:      29

Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 50 summaries

Database :        SwissProt_39:*
```

Minimum DB seq length: 0
Maximum DB seq length: 5

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	DB ID	Description
1	8	44.4	5	1	AL14_CARMA	P81817	carcinus ma
2	7	38.9	4	1	FAR3_HIRME	P42562	hirudo medi
3	7	38.9	4	1	FAR4_HIRME	P42563	hirudo medi
4	7	38.9	4	1	FARP_ARTTR	P41853	artioposthi
5	7	38.9	4	1	FARP_PERAM	P01373	periplaneta
6	7	38.9	4	1	PSK_DAUCA	P58261	daucus caro
7	5	27.8	4	1	FLRF_HIRME	P42561	hirudo medi
8	5	27.8	4	1	FMRF_MACNI	P01162	macrocallis
9	5	27.8	4	1	TUFT_HUMAN	P01858	homo sapien
10	5	27.8	5	1	BIOB_SALTY	P12678	salmonella
11	5	27.8	5	1	UFO1_MOUSE	P38639	mus musculu
12	4	22.2	3	1	GRWM_HUMAN	P01157	homo sapien
13	3	16.7	4	1	ACH1_ACHFU	P35904	achatina fu
14	3	16.7	4	1	ACH1_ACHFU	P81864	pardachirus
15	3	16.7	5	1	SUGA_ACHDO	P19991	acheta dome
16	3	16.7	5	1	TPIS_CANFA	P54714	canis famil
17	3	16.7	5	1	UC22_MAIZE	P80628	zea mays (m
18	3	16.7	3	1	LUXE_VIBFI	P24472	vibrio fisc
19	2	11.1	3	1	THYL_PIG	P01151	sus scrofa
20	2	11.1	4	1	DCML_PSECH	P01956	pseudomonas
21	2	11.1	4	1	DCMS_PSECH	P19916	pseudomonas
22	2	11.1	5	1	BIOA_CITFR	P19918	citrobacter
23	2	11.1	5	1	BIOA_CITFR	P13071	citrobacter
24	2	11.1	5	1	BIOA_SALTY	P12297	salmonella
25	2	11.1	5	1	BIOB_CITFR	P30425	citrobacter
26	2	11.1	5	1	BPP7_BOTIN	P13973	bothrops in
27	1	5.6	4	1	TRM3_ECOLI	P02731	escherichia
28	1	5.6	4	1	EOS1_HUMAN	P36515	homo sapien
29	0	0.0	5	1	RM01_YEAST	P38005	saccharomyc
					UXA4_CHLTR		chlamydia t

ALIGNMENTS

```
RESULT    1
AL14_CARMA
ID   AL14_CARMA           STANDARD;      PRT;        5 AA.
AC   P81817;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 14.
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5      5       AMIDATION (POTENTIAL).
SQ   SEQUENCE    5 AA;    586 MW;  67287905AB300000 CRC64;

Query Match         44.4%;  Score 8;  DB 1;  Length 5;
Best Local Similarity   40.0%;  Pred. No. 1e+05;
Matches    2;  Conservative   0;  Mismatches   3;  Indels   0;  Gaps   0;

Qy        1 YRXXL 5
            |   |
Db        1 YSFGL 5


RESULT    2
FAR3_HIRME
ID   FAR3_HIRME           STANDARD;      PRT;        4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE YLRF-AMIDE.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE    4 AA;    598 MW;  69D4073B0000000 CRC64;

Query Match         38.9%;  Score 7;  DB 1;  Length 4;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    1;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy        1 Y 1
            |
Db        1 Y 1
```

RESULT 3

```
ID   FAR4_HIRME          STANDARD;      PRT;      4 AA.
AC   P42563;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE YMRF-AMIDE.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Life Sci. 17:1253-1256(1975).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       4      4       AMIDATION.
SQ   SEQUENCE   4 AA;    616 MW;  69D4066B30000000 CRC64;
```

```
Qy   1 y 1
     |  |
Db   1 y 1
```

```
Query Match             38.9%;  Score 7;  DB 1;  Length 4;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
     Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

RESULT 4

```
ID   FARP_ARTTR          STANDARD;      PRT;      5 AA.
AC   P41853;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE RYIRF-AMIDE.
OS   Artioposthia triangulata.
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellaria Platyhelminths;
OC   Rhabditophora; Seriata; Tricladida; Terricola; Geoplanidae;
OC   Arthurdendyus.
OX   NCBI_TaxID=132421;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=94211927; PubMed=7909164;
RA   Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT   "RYIRFamide: a turbellarian FMRFamide-related peptide (FaRP).";
RL   Regul. Pept. 50:37-43(1994).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;    754 MW;  69D4004B44600000 CRC64;
```

```
Qy   1 y 1
     |  |
Db   2 y 2
```

```
Query Match             38.9%;  Score 7;  DB 1;  Length 5;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
     Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

RESULT 5

```
ID   PRCT_PERAM          STANDARD;      PRT;      5 AA.
AC   PC1373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   PROCTOLIN.
OS   Periplaneta americana (American cockroach),
OS   Limulus polyphemus (Atlantic horseshoe crab), and
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattodea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 6850, 6759;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P. americana;
RX   MEDLINE=76074708; PubMed=576;
RA   Starratt A.N., Brown B.E.;
RT   "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT   in insects.";
RL   Life Sci. 17:1253-1256(1975).
RN   [2]
RP   BIOLOGICAL SOURCE.
RC   SPECIES=P.americana;
RX   MEDLINE=81225865; PubMed=6113690;
RA   O'Shea M., Adams M.E.;
RT   "Pentapeptide (proctolin) associated with an identified neuron.";
RL   Science 213:567-569(1981).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L. polyphemus;
RX   MEDLINE=90287800; PubMed=2356151;
RA   Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA   Shabanowitz J.;
RT   "Identification of proctolin in the central nervous system of the
RT   horseshoe crab, Limulus polyphemus.";
RL   Peptides 11:205-211(1990).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=C.maenas;
RX   MEDLINE=86232789; PubMed=2872661;
RA   Stangier J., Dircksen H., Keller R.;
RT   "Identification and immunocytochemical localization of proctolin in
RT   pericardial organs of the shore crab, Carcinus maenas.";
RL   Peptides 7:67-72(1986).
CC   -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,
CC       MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC       THE CRAB PERICARDIAL ORGANS.
DR   PIR; A01644; HOROHA.
DR   PIR; A60411; A60411.
KW   Neuropeptide.
SQ   SEQUENCE   5 AA;    649 MW;  71B76733B44600000 CRC64;
```

```
Qy   1 y 1
     |  |
Db   2 y 2
```

```
Query Match             38.9%;  Score 7;  DB 1;  Length 5;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
     Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

RESULT 6

```
ID   PSK_DAUCA           STANDARD;      PRT;      5 AA.
AC   P58261;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PHYTOSULFOKINE-ALPHA (PSK-ALPHA) [CONTAINS: PHYTOSULFOKINE-BETA (PSK-
DE   BETA)].
OS   Daucus carota (Carrot).
```

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:39:26 ; Search time 25.34 Seconds
                                            (without alignments)
                                            15.030 Million cell updates/sec

Title: US-09-281-760C-2
Perfect score: 18
Sequence: 1 YRXXL 5
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 205

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   PIR_68:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Minimum DB seq length: 0
Maximum DB seq length: 5

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	12	66.7	3	2	A22565	R-phycoerythrin al
2	12	66.7	3	2	F22565	R-phycoerythrin ga
3	8	44.4	5	2	I40469	dn2X-like protein
4	7	38.9	3	3	A43391	TRH-like tripeptid
5	7	38.9	3	2	S68328	blood cell protein
6	7	38.9	4	2	A32039	tyrosine-melanocyt
7	7	38.9	4	2	A37832	phenol 2-monooxyge
8	7	38.9	4	2	S09478	globulin IV alpha
9	7	38.9	4	2	I61883	protamine P1 - ora
10	7	38.9	4	2	PT0240	protamine P1 - ora
11	7	38.9	4	2	S43959	Ig heavy chain CRD
12	7	38.9	4	2	I37013	Ig mu chain V regi
13	7	38.9	4	2	I84439	protamine P1 - Cer
14	7	38.9	5	1	HOROHA	protamine P1 - sav
15	7	38.9	4	2	B37325	pap fimbrial regul
16	7	38.9	4	2	I39964	proctolin - Americ
17	7	38.9	4	2	I39966	ribosomal protein
18	7	38.9	4	2	I39965	ribosomal protein
19	7	38.9	4	2	E60274	ribosomal protein
20	7	38.9	4	2	Q0009	major protein anti
21	7	38.9	4	2	PQ0689	angiotensin-conver
22	7	38.9	4	2	B61445	photosystem I 10.4
23	7	38.9	4	2	A61445	Leu-enkephalin - b
24	7	38.9	4	2	A60411	Met-enkephalin - b
25	7	38.9	4	2	PT0278	proctolin - Atlant
26	7	38.9	5	2	S68326	Ig heavy chain CRD
27	7	27.8	3	2	PQ0010	blood cell protein
28	5	27.8	4	1	ECXAA	angiotensin-conver
29	5	27.8	4	2	A02147	antho-RFamide neur
30	5	27.8	4	2	ECNK	phagocytosis-stimu
31	5	27.8	4	2	D41654	cardioexcitatory n
32	5	27.8	4	2	I40870	hypothetical prote
33	5	27.8	4	2	JQ1273	phospholipase C (E
34	5	27.8	4	2	A35779	neuropeptide Antho
35	5	27.8	4	2	A25844	neuropeptide Antho
36	5	27.8	4	2	A60418	autho-RF amide neu
37	5	27.8	4	2	PT0721	FMRFamide - polych
38	5	27.8	4	2	PT0721	T-cell receptor be
39	5	27.8	4	2	S47552	ubiquitin - rat
40	5	27.8	4	2	JN0862	peptidyl-dipeptida
41	5	27.8	4	2	I40702	primase - Citrobac
42	5	27.8	4	2	A44955	alkanal monooxygen
43	5	27.8	4	2	D60274	major protein anti
44	5	27.8	4	2	B22565	R-phycoerythrin al
45	5	27.8	4	2	T14910	hypothetical prote
46	5	27.8	4	2	A60803	neuropeptide - sea
47	5	27.8	4	2	S53595	hypothetical prote
48	5	27.8	4	2	PT0295	Ig heavy chain CRD
49	5	27.8	4	2	S62883	seminal plasma pro
50	5	27.8	4	2	PT0513	T-cell receptor be
					PT0525	T-cell receptor be

ALIGNMENTS

RESULT   1
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: A22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KLO>

Query Match   66.7%;   Score 12;   DB 2;   Length 3;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches 2;   Conservative 0;   Mismatches 0;   Indels. 0;   Gaps 0;

Qy   1 YR 2
           ||
Db   2 YR 3

RESULT   2
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: F22565; MUID:85182601
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

Query Match   66.7%;   Score 12;   DB 2;   Length 5;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches 2;   Conservative 0;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   1 YR 2
           ||
Db   4 YR 5

RESULT 3
I40469
dnaX-like protein - Bacillus subtilis (fragment)
C;Species: Bacillus subtilis
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 21-Jul-2000
C;Accession: I40469
R;Struck, J.C.; Hartmann, R.K.; Toschka, H.Y.; Erdmann, V.A.
R;Mol. Gen. Genet. 215, 478-482, 1989
A;Title: Transcription and Processing of Bacillus subtilis small cytoplasmic RNA.
A;Reference number: I40469; MUID:89218958
A;Accession: I40469
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: EMBL:X14796; NID:g40130; PIDN:CAA32902.1; PID:g4376204
C;Genetics:
A;Start codon: GTG

Qy   1 YR 2
         |:|
Db   3 YQ 4

Query Match   44.4%;   Score 8;   DB 2;   Length 5;
Best Local Similarity   50.0%;   Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   0;   Indels   0;   Gaps   0;

RESULT 4
A4391
TRH-like tripeptide - alfalfa
C;Species: Medicago sativa (alfalfa)
C;Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 07-May-1999
C;Accession: A43391
R;Lackey, D.B.
J. Biol. Chem. 267, 17508-17511, 1992
A;Title: Isolation and structural determination of a novel TRH-like tripeptide, pyroGlu-
A;Reference number: A43391; MUID:92388092
A;Accession: A43391
A;Molecule type: protein
A;Residues: 1-3 <LAC>
A;Keywords: amidated carboxyl end; pyroglutamic acid
F;1/Modified site: pyrrolidone carboxylic acid (Gln) #status experimental
F;3/Modified site: amidated carboxyl end (Pro) #status experimental

Qy   1 Y 1
         |
Db   2 Y 2

Query Match   38.9%;   Score 7;   DB 2;   Length 3;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches   1; Conservative   0; Mismatches   0;   Indels   0;   Gaps   0;

RESULT 5
S68328
blood cell protein A - Molgula manhattensis (fragment)
C;Species: Molgula manhattensis
C;Date: 19-Mar-1997 #sequence_revision 29-Aug-1997 #text_change 11-Jun-1999
C;Accession: S68328
R;Taylor, S.W.; Ross, M.M.; Waite, J.H.
Arch. Biochem. Biophys. 324, 228-240, 1995
A;Title: Novel 3,4-di- and 3,4,5-trihydroxyphenylalanine-containing polypeptides from th
A;Reference number: S68325; MUID:96132650
A;Accession: S68328
A;Molecule type: protein
A;Residues: 1-3 <TAY>

Query Match   38.9%;   Score 7;   DB 2;   Length 3;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;

Matches   1; Conservative   0; Mismatches   0;   Indels   0;   Gaps   0;

RESULT 6
A32039
tyrosine-melanocyte-stimulating hormone release-inhibiting factor 1 - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 31-Jul-1989 #sequence_revision 31-Jul-1989 #text_change 18-Aug-2000
C;Accession: A32039
R;Horvath, A.; Kastin, A.J.
J. Biol. Chem. 264, 2175-2179, 1989
A;Title: Isolation of tyrosine-melanocyte-stimulating hormone release-inhibiting fact
A;Reference number: A32039; MUID:89123285
A;Accession: A32039
A;Molecule type: protein
A;Residues: 1-4 <HOR>
A;Experimental source: brain
A;Superfamily: unassigned animal peptides
C;Keywords: amidated carboxyl end
F;4/Modified site: amidated carboxyl end (Gly) #status experimental

Qy   1 Y 1
         |
Db   1 Y 1

Query Match   38.9%;   Score 7;   DB 2;   Length 4;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches   1; Conservative   0; Mismatches   0;   Indels   0;   Gaps   0;

RESULT 7
A37832
phenol 2-monooxygenase (EC 1.14.13.7) chain P5 - Pseudomonas sp. (strain CF600) (frag
C;Species: Pseudomonas sp.
C;Date: 14-Jun-1991 #sequence_revision 14-Jun-1991 #text_change 23-Jun-1993
C;Accession: A37832
R;Powlowski, J.; Shingler, V.
J. Bacteriol. 172, 6834-6840, 1990
A;Title: In vitro analysis of polypeptide requirements of multicomponent phenol hydro
A;Reference number: A37832; MUID:91072231
A;Accession: A37832
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-4 <POW>
C;Keywords: oxidoreductase

Qy   1 Y 1
         |
Db   2 Y 2

Query Match   38.9%;   Score 7;   DB 2;   Length 4;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches   1; Conservative   0; Mismatches   0;   Indels   0;   Gaps   0;

RESULT 8
S09478
globulin IV alpha subunit gamma chain, seed - cucurbit (fragment)
N;Alternate names: 11S globulin alpha subunit gamma chain
C;Species: Cucurbita sp. (cucurbit)
C;Date: 21-Nov-1993 #sequence_revision 08-Nov-1996 #text_change 08-Nov-1996
C;Accession: S09478
R;Ohmiya, M.; Hara, I.; Matsubara, H.
Plant Cell Physiol. 21, 157-167, 1980
A;Title: Pumpkin (Cucurbita sp.) seed globulin IV. Terminal sequences of the acidic a
A;Reference number: S09066
A;Accession: S09478

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:19:20 ; Search time 23.48 Seconds
                                             (without alignments)
                                             4.792 Million cell updates/sec

Title: US-09-281-760C-1
Perfect score: 18
Sequence: 1 LXXYR 5
Scoring table: BLOSUM62
                   Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 16280
inimum DB seq length: 0
aximum DB seq length: 5

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   Issued_Patents_AA:*
          1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
          2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
          3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
          4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
          5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
          6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	% Query Match	Match Length	DB	DB ID	Description
1	66.7	2	4		US-09-214-614-2	Sequence 2, Appl1
2	66.7	3	2		US-08-968-676-27	Sequence 27, Appl
3	66.7	2	2		US-08-706-178-2	Sequence 4, Appli
4	66.7	4	1		US-08-240-712-4	Sequence 4, Appli
5	66.7	4	1		US-08-428-488-59	Sequence 59, Appl
6	66.7	4	1		US-08-428-488-60	Sequence 60, Appl
7	66.7	4	1		US-08-428-488-61	Sequence 61, Appl
8	66.7	4	1		US-08-428-488-62	Sequence 62, Appl
9	66.7	4	1		US-08-428-488-63	Sequence 63, Appl
10	66.7	4	1		US-08-428-488-64	Sequence 64, Appl
11	66.7	4	1		US-08-428-488-65	Sequence 65, Appl
12	66.7	4	1		US-08-428-488-66	Sequence 66, Appl
13	66.7	4	1		US-08-428-488-67	Sequence 67, Appl
14	66.7	4	1		US-08-428-488-68	Sequence 68, Appl
15	66.7	4	1		US-08-428-488-69	Sequence 69, Appl
16	66.7	4	1		US-08-428-488-70	Sequence 70, Appl
17	66.7	4	1		US-08-428-488-71	Sequence 71, Appl
18	66.7	4	1		US-08-428-488-72	Sequence 72, Appl
19	66.7	4	1		US-08-428-488-73	Sequence 73, Appl
20	66.7	4	1		US-08-428-488-74	Sequence 74, Appl
21	66.7	4	1		US-08-255-272-22	Sequence 22, Appl
22	66.7	4	1		US-08-335-198-25	Sequence 25, Appl
23	66.7	4	1		US-08-429-732-20	Sequence 20, Appl
24	66.7	4	1		US-08-203-806B-15	Sequence 15, Appl
25	66.7	4	1		US-08-443-890-4	Sequence 4, Appli
26	66.7	4	2		US-08-441-871-1	Sequence 1, Appli
27	66.7	4	2		US-08-441-871-63	Sequence 63, Appl
28	66.7	4	2		US-08-441-871-65	Sequence 65, Appl
29	66.7	4	2		US-08-441-871-66	Sequence 66, Appl
30	66.7	4	2		US-08-441-871-68	Sequence 68, Appl
31	66.7	4	2		US-08-441-871-70	Sequence 70, Appl
32	66.7	4	2		US-08-473-022-2	Sequence 2, Appl1
33	66.7	4	2		US-08-705-660-27	Sequence 27, Appl
34	66.7	4	2		US-08-705-660-28	Sequence 28, Appl
35	66.7	4	2		US-08-623-933B-61	Sequence 61, Appl
36	66.7	4	2		US-08-685-589A-6	Sequence 6, Appli
37	66.7	4	2		US-08-685-589A-62	Sequence 62, Appl
38	66.7	4	2		US-08-968-676-26	Sequence 26, Appl
39	66.7	4	2		US-08-752-844-12	Sequence 12, Appl
40	66.7	4	2		US-08-747-137-78	Sequence 78, Appl
41	66.7	4	2		US-08-591-196-12	Sequence 12, Appl
42	66.7	4	2		US-08-989-045-27	Sequence 27, Appl
43	66.7	4	3		US-08-989-045-28	Sequence 28, Appl
44	66.7	4	3		US-08-986B-20	Sequence 20, Appl
45	66.7	4	4		US-09-082-593-11	Sequence 11, Appl
46	66.7	4	4		US-09-082-593-11	Sequence 4, Appli
47	66.7	4	5		US-09-058-562-4	Sequence 62, Appl
48	66.7	4	5		PCT-US92-09752-4	Sequence 4, Appli
49	66.7	4	5		PCT-US93-01669-16	Sequence 16, Appl
50	66.7	4	6		PCT-US96-08798-2	Sequence 2, Appli
					5312899-1	Patent No. 5312899

ALIGNMENTS

RESULT 1
US-09-214-614-2
Sequence 2, Application US/09214614
Patent No. 6225100
GENERAL INFORMATION:
APPLICANT: Grund, Alan D.
APPLICANT: Maurina-Brunker, Julie
TITLE OF INVENTION: NOVEL ARYLSULFOTRANSFERASE
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sheridan Ross P.C.
STREET: 1700 Lincoln Street, Suite 3500
CITY: Denver
STATE: Colorado
COUNTRY: U.S.A
ZIP: 80203
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/214,614
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Connell, Gary J.
REGISTRATION NUMBER: 32,020
REFERENCE/DOCKET NUMBER: 3161-15-PCT
TELECOMMUNICATION INFORMATION:
TELEPHONE: 303/863-9700
TELEFAX: 303/863-0223
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-214-614-2

Query Match   66.7%;   Score 12;   DB 4;   Length 2;
Best Local Similarity   100.0%;   Pred. No. 1.6e+05;
Matches   2; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 2
US-08-968-676-27
Sequence 27, Application US/08968676
Patent No. 5919639
GENERAL INFORMATION:
APPLICANT: Humphreys, Robert E
APPLICANT: Adams, Sharlene
APPLICANT: Xu, Minzhen
TITLE OF INVENTION: IMMUNOTHERAPY BY MODULATION OF ANTIGEN
NUMBER OF SEQUENCES: 165
CURRENT APPLICATION DATA:
CURRENT APPLICATION NUMBER: US/08,968,676
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kevin M. Farrell, P.C.
STREET: P. O. Box 999
CITY: York Harbor
STATE: ME
COUNTRY: USA
ZIP: 03911
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08,968,676
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Farrell, Kevin M
REGISTRATION NUMBER: 35,505
REFERENCE/DOCKET NUMBER: REH-9601
TELECOMMUNICATION INFORMATION:
TELEPHONE: (207) 363-0528
TELEFAX: (207) 363-0558
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 3 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-968-676-27

Query Match 66.7%; Score 12; DB 2; Length 3;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YR 5
|| ||
Db 1 YR 2

RESULT 3
US-08-240-712-4
Sequence 4, Application US/08240712
Patent No. 5559907
GENERAL INFORMATION:
APPLICANT: ANDERSON, DAVID C.
APPLICANT: MATHEWS, ANTONY JAMES
APPLICANT: STETLER, GARY L.
TITLE OF INVENTION: PRODUCTION AND USE OF MULTIMERIC
TITLE OF INVENTION: HEMOGLOBINS
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE; D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/240,712
FILING DATE: 09-MAY-1994
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/09752
FILING DATE: 13-MAY-1993
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: ANDERSON-6
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 4 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-240-712-4

Query Match 66.7%; Score 12; DB 1; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YR 5
|| ||
Db 3 YR 4

RESULT 4
US-08-428-488-59
Sequence 59, Application US/08428488
Patent No. 5624894
GENERAL INFORMATION:
APPLICANT: BODOR, Nicholas S.
TITLE OF INVENTION: BRAIN-ENHANCED DELIVERY OF NEUROACTIVE
TITLE OF INVENTION: PEPTIDES BY SEQUENTIAL METABOLISM
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burns, Doane, Swecker & Mathis
STREET: P.O. Box 1404
CITY: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22313-1404
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/428,488
FILING DATE: 27-APR-1995
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Baumeister, Mary Katherine
REGISTRATION NUMBER: 26,254
REFERENCE/DOCKET NUMBER: 028724-087
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021

```
OM protein - protein search, using sw model

                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

Run on:             April 21, 2002, 10:18:10 ; Search time 40.45 Seconds
                    (without alignments)
                    9.156 Million cell updates/sec

Title:              US-09-281-760C-1
Perfect score:      18
Sequence:           1 LXXYR 5
Scoring table:      BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched:           522463 seqs, 74072290 residues

Total number of hits satisfying chosen parameters:      20141

Post-processing: Minimum Match 0%
●Minimum DB seq length: 0
 Maximum Match 100%
 ximum DB seq length: 5
                    Listing first 50 summaries

Database :          A_Geneseq_1101:*

           1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
           2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
           3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
           4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
           5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
           6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
           7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
           8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
           9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
          10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
          11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
          12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
          13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
          14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
          15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
          16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
          17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
          18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
          19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
          20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
          21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
          22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*
```

SUMMARIES

Result No.	% Score	Query Match Length	DB	DB ID	Description
1	77.8	5	21	AAY69581	Generic immunogeni
2	66.7	2	22	AAB91738	Opioid peptide SEQ
3	66.7	3	18	AAW33241	Analgesic peptide.
4	66.7	3	19	AAW56224	Anti-inflammatory
5	66.7	3	22	AAB91597	Opioid peptide SEQ
6	66.7	4	5	AAP40467	Enkephalin analogu
7	66.7	4	6	AAP50336	Analgesic enkephal
8	66.7	4	6	AAP50690	Analgesic tetrapep
9	66.7	4	6	AAP61491	Analgesic peptide.
10	66.7	4	7	AAP61741	Analgesic peptide.
11	66.7	4	8	AAP71341	Analgesic peptide

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

```
12  12  66.7  4   8   AAP71580  Growth hormone rel
13  12  66.7  4   8   AAP71582  Growth hormone rel
14  12  66.7  4   8   AAP71700  Analgesic peptide
15  12  66.7  4   13  AAR23084  Fluorinated peptid
16  12  66.7  4   13  AAR24410  CPase B-like enzym
17  12  66.7  4   14  AAR41643  Internalisation si
18  12  66.7  4   14  AAR33660  Growth hormone rel
19  12  66.7  4   14  AAR33664  Growth hormone rel
20  12  66.7  4   14  AAR38136  Protease-non labil
21  12  66.7  4   15  AAW51593  Protease-non labil
22  12  66.7  4   15  AAW51589  Peptide analogue #
23  12  66.7  4   15  AAW51595  Peptide analogue #
24  12  66.7  4   15  AAW51597  Peptide analogue #
25  12  66.7  4   15  AAR55498  Opioid oligopeptid
26  12  66.7  4   15  AAR55497  Opioid oligopeptid
27  12  66.7  4   15  AAR55502  Opioid oligopeptid
28  12  66.7  4   15  AAR55503  Opioid oligopeptid
29  12  66.7  4   15  AAR55507  Opioid oligopeptid
30  12  66.7  4   15  AAR55506  Opioid oligopeptid
31  12  66.7  4   15  AAR55481  Opioid oligopeptid
32  12  66.7  4   15  AAR55484  Opioid oligopeptid
33  12  66.7  4   15  AAR55485  Opioid oligopeptid
34  12  66.7  4   15  AAR55488  Opioid oligopeptid
35  12  66.7  4   15  AAR55489  Opioid oligopeptid
36  12  66.7  4   15  AAR55493  Opioid oligopeptid
37  12  66.7  4   15  AAR55494  Opioid oligopeptid
38  12  66.7  4   15  AAR55465  Opioid oligopeptid
39  12  66.7  4   15  AAR55467  Opioid oligopeptid
40  12  66.7  4   15  AAR55469  Opioid oligopeptid
41  12  66.7  4   15  AAR55471  Opioid oligopeptid
42  12  66.7  4   15  AAR55473  Opioid oligopeptid
43  12  66.7  4   15  AAR55475  Opioid oligopeptid
44  12  66.7  4   15  AAR55477  Opioid oligopeptid
45  12  66.7  4   15  AAR55479  Opioid oligopeptid
46  12  66.7  4   15  AAR55449  Opioid oligopeptid
47  12  66.7  4   15  AAR55451  Opioid oligopeptid
48  12  66.7  4   15  AAR55453  Opioid oligopeptid
49  12  66.7  4   15  AAR55455  Opioid oligopeptid
50  12  66.7  4   15  AAR55435  Opioid oligopeptid
```

ALIGNMENTS

```
RESULT      1
AAY69581

ID    AAY69581 standard; peptide; 5 AA.
XX
AC    AAY69581;
XX
DT    19-APR-2000  (first entry)
XX
DE    Generic immunogenic peptide SEQ ID NO:6, for anti-IgE Ab generation.
XX
KW    Immunogenic; autoimmunity; IgE autoantibody; immunoglobulin E;
KW    allergy; canine; prophylaxis; treatment.
OS    Synthetic.
XX
FH    Key            Location/Qualifiers
FT
FT    Misc-difference 2..3
FT                    /label= Xaa
FT                    /note= "Xaa is any amino acid"
XX
PN    EP955311-A2.
XX
PD    10-NOV-1999.
XX
PF    09-APR-1999; 99EP-0107020.
XX
PR    09-APR-1998; 98US-0058332.
PR    30-MAR-1999; 99US-0281761.
XX
```

PA (IDEX-) IDEXX LAB INC.
XX
PI Lawton R, Mermer B, Francoeur G;
XX
DR WPI; 2000-111182/11.
XX
PT New peptides, useful for generating canine autoantibodies against
PT immunoglobulin E (IgE) to treat allergic symptoms -
PS Claim 2; Page 5; 38pp; English.
XX
CC Sequences AAY69576-Y69586 and AAY69598-Y69599 represent peptides that may
CC be used to induce anti-IgE (immunoglobulin E) autoantibodies in a dog.
CC IgE is an important mediator of canine allergic responses, including Type
CC I immediate hypersensitivity. It is estimated that up to 30% of all dogs
CC suffer from allergies or allergy-related skin disorders, and allergic
CC dermatitis is thought to affect between 3 and 15% of dogs. Common canine
CC allergens include fleas, pollens, molds and dust, with flea allergy
CC being the most frequent. Once IgE has been released in response to an
CC allergen, the IgE binds to receptors on mast cells. The IgE molecules
CC can then become cross-linked, which induces the mast cells to release
CC histamine, which leads to the manifestation of allergic symptoms. The
CC peptides of the invention correspond to epitopes of canine IgE, or
CC conservative variations thereof, that are accessible in circulating IgE
CC but not in IgE bound to mast cells. The autoantibodies produced on
CC immunisation with such peptides therefore target only circulating IgE,
CC avoiding the possibility that they may cross-link bound IgE molecules,
CC which would exacerbate an ongoing allergic response even in the absence
CC of allergen. Anti-IgE autoantibody induction results in an immune
CC response that targets the IgE-expressing B cell, which affects IgE
CC synthesis by downregulating IgE production by B cells and/or targeting
CC the IgE-producing B cell for destruction. The peptides of the invention
CC are therefore useful for the treatment or prophylaxis of allergic
CC symptoms in a dog.
XX
SQ Sequence 5 AA;
XX

Query Match   77.8%; Score 14; DB 21; Length 5;
Best Local Similarity   100.0%; Pred. No. 4.3e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   1 LXXYR 5
          |||||
Db   1 lxxyr 5

RESULT 2
AAB91738
ID AAB91738 standard; Peptide; 2 AA.
XX
AC AAB91738;
XX
DT 22-JUN-2001 (first entry)
XX
DE Opioid peptide SEQ ID NO:914.
XX
OS Homo sapiens.
OS Synthetic.
XX
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification; succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
PN WO200069900-A2.
XX
PD 23-NOV-2000.
XX
PF 17-MAY-2000; 2000WO-US13576.
XX
PR 17-MAY-1999; 99US-0134406.
PR 10-SEP-1999; 99US-0153406.
PR 15-OCT-1999; 99US-0159783.

XX
PA (CONJ-) CONJUCHEM INC.
XX
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudeau K;
XX
DR WPI; 2001-112059/12.
XX
PT Modifying and attaching therapeutic peptides to albumin prevents
PT peptidase degradation, useful for increasing length of in vivo activity
PT -
PS Disclosure; Page 492; 733pp; English.
XX
CC The present invention describes a modified therapeutic peptide (I)
CC comprising a therapeutically active amino acid region (III) and a
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to
CC a less therapeutically active amino acid region (IV), which covalently
CC bonds with amino/hydroxyl/thiol groups on blood components to form a
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth
CC factors and neurotransmitters, to protect them from peptidase activity
CC in vivo for the treatment of various disorders. Endogenous therapeutic
CC peptides are not suitable as drug candidates as they require frequent
CC administration due to rapid degradation by peptidases in the body.
CC Modifying and attaching therapeutic peptides to albumin prevents or
CC reduces the action of peptidases to increase length of activity (half
CC life) and specificity as bonding to large molecules decreases
CC intracellular uptake and interference with physiological processes.
CC AAB90829 to AAB92441 represent peptides which can be used in the
CC exemplification of the present invention.
XX
SQ Sequence 2 AA;
XX

Query Match   66.7%; Score 12; DB 22; Length 2;
Best Local Similarity   100.0%; Pred. No. 4.3e+05;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   4 YR 5
          ||
Db   1 yr 2

RESULT 3
AAW33241
ID AAW33241 standard; peptide; 3 AA.
XX
AC AAW33241;
XX
DT 30-JAN-1998 (first entry)
XX
DE Analgesic peptide.
XX
KW Opioid receptor; treatment; prevention; pain; cancer; analgesic.
XX
OS Synthetic.
XX
FH Key   Location/Qualifiers
FT Modified-site   1
FT   /note= "HN=C(CH3)-Tyr"
FT Misc-difference 2
FT   /note= "D-form residue"
FT Modified-site   3
FT   /note= "Phe-N(CH3)-CH2-CH2-COOH"
XX
PN WO9710261-A1.
XX
PD 20-MAR-1997.
XX
PF 10-SEP-1996; 96WO-JP02572.
XX
PR 11-SEP-1995; 95JP-0231161.

```
                    GenCore version 4.5
                    Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:            April 21, 2002, 10:38:16 ; Search time 37.83 Seconds
                                              (without alignments)
                                              19.333 Million cell updates/sec

Title:             US-09-281-760C-1
Perfect score:     18
Sequence:          1 LXXYR 5

Scoring table:     BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched:          473505 seqs, 14627 2329 residues

Total number of hits satisfying chosen parameters:       8

        imum DB seq length: 0
        imum DB seq length: 5

Post-processing:   Minimum Match 0%
                   Maximum Match 100%
                   Listing first 50 summaries

Database :         SPTREMBL_17:*
                   1: sp_archea:*
                   2: sp_bacteria:*
                   3: sp_fungi:*
                   4: sp_human:*
                   5: sp_invertebrate:*
                   6: sp_mammal:*
                   7: sp_mhc:*
                   8: sp_organelle:*
                   9: sp_phage:*
                   10: sp_plant:*
                   11: sp_rodent:*
                   12: sp_virus:*
                   13: sp_vertebrate:*
                   14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB ID	ID	Description
1	22.2	4	11		Q08433	rattus norv
2	16.7	5	13		P82070	litoria rub
3	16.7	5	13		P82071	litoria rub
4	16.7	5	13		P82072	litoria rub
5	16.7	5	13		P82073	litoria rub
6	16.7	5	13		P82099	litoria rub
7	16.7	5	13		P82100	litoria rub
8	11.1	5	10		Q99007	hordeum vul

ALIGNMENTS

```
RESULT 1
ID   Q08433
AC   Q08433;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JAN-1999 (TrEMBLrel. 09, Last annotation update)
DE   UDP-GLUCURONOSYLTRANSFERASE, MICROSOMAL (EC 2.4.1.17) (UDPGT)
DE   (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GUNN;
RX   MEDLINE=91282758; PubMed=1840486;
RA   Sato H., Aono S., Kashiwamata S., Koiwai O.;
RT   "Genetic defect of bilirubin UDP-glucuronosyltransferase in the
RT   hyperbilirubinemic Gunn rat.";
RL   Biochem. Biophys. Res. Commun. 177:1161-1164(1991).
CC   -!- FUNCTION: UDPGT IS OF MAJOR IMPORTANCE IN THE CONJUGATION AND
CC       SUBSEQUENT ELIMINATION OF POTENTIALLY TOXIC XENOBIOTICS AND
CC       ENDOGENOUS COMPOUNDS.
CC   -!- CATALYTIC ACTIVITY: UDP-GLUCURONATE + ACCEPTOR = UDP + ACCEPTOR
CC       BETA-D-GLUCURONOSIDE.
CC   -!- SUBCELLULAR LOCATION: MICROSOME.
DR   EMBL; S38636; AAB19259.1; -.
KW   Transferase; Glycosyltransferase; Microsome; Multigene family.
FT   NON_TER        1        1
FT   NON_TER        4        4
SQ   SEQUENCE       4 AA;    473 MW;   633732C420000000 CRC64;

Query Match             22.2%;    Score 4;   DB 11;   Length 4;
Best Local Similarity  100.0%;    Pred. No. 4.7e+05;
Matches     1;  Conservative    0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      1 L 1
          |
Db      3 L 3


RESULT 2
P82070
ID   P82070
AC   P82070;
DT   01-MAY-2000  (TrEMBLrel. 13, Created)
DT   01-MAY-2000  (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000  (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 1.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella', the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=598; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE    5 AA;    598 MW;  6DD9C9CAB2A00000 CRC64;

PRELIMINARY;        PRT;        4 AA.

PRELIMINARY;        PRT;        5 AA.

Query Match             16.7%;    Score 3;   DB 13;   Length 5;
Best Local Similarity    0.0%;    Pred. No. 4.7e+05;
Matches     0;  Conservative    1;  Mismatches   0;  Indels   0;  Gaps   0;

Qy      4 Y 4
          :
```

```
RESULT 3

ID   P82071        PRELIMINARY;      PRT;    5 AA.
AC   P82071;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 2.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella'. the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW   Amphibian skin; Amidation.
SQ   SEQUENCE   5 AA;  626 MW;  6DD9C9CB10300000 CRC64;

Query Match             16.7%;  Score 3;  DB 13;  Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
  Matches    0;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 Y 4
Db          3 F 3


RESULT 4

ID   P82072        PRELIMINARY;      PRT;    5 AA.
AC   P82072;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 3.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella'. the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian skin; Amidation.
SQ   SEQUENCE   5 AA;  656 MW;  71A9C9CB10300000 CRC64;

Query Match             16.7%;  Score 3;  DB 13;  Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
  Matches    0;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 Y 4
Db          3 F 3


RESULT 5

ID   P82073        PRELIMINARY;      PRT;    5 AA.
AC   P82073;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 3.2.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN SECRETION;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litoria electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:0-0(1999).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
KW   Amphibian skin; Amidation.
SQ   SEQUENCE   5 AA;  570 MW;  71A9C9C862A00000 CRC64;

Query Match             16.7%;  Score 3;  DB 13;  Length 5;
Best Local Similarity   0.0%;   Pred. No. 4.7e+05;
  Matches    0;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          4 Y 4
Db          3 F 3


RESULT 6

ID   P82099        PRELIMINARY;      PRT;    5 AA.
AC   P82099;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   ELECTRIN 3.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=SKIN SECRETION;
RA   Wabnitz P.A., Bowie J.H., Tyler M.J., Wallace J.C.;
RT   "Peptides from the skin glands of the Australian buzzing tree frog
RT   Litori electrica. Comparison with the skin peptides from Litoria
RT   rubella.";
RL   Aust. J. Chem. 52:0-0(1999).
KW   Amphibian skin; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;  630 MW;  668761F2C9A00000 CRC64;

Query Match             16.7%;  Score 3;  DB 13;  Length 5;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:38:51 ; Search time 17.3 Seconds
(without alignments)
10.597 Million cell updates/sec

Title: US-09-281-760C-1
Perfect score: 18
Sequence: 1 LXXYR 5

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 29

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

nimum DB seq length: 0
ximum DB seq length: 5

Database :   SwissProt_39:*

* Pred. No. is the number of results predicted by chance to have a
  score greater than or equal to the score of the result being printed,
  and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	7	38.9	4	1	FAR3_HIRME	P42562; hirudo medi
2	7	38.9	4	1	FAR4_HIRME	P42563; hirudo medi
3	7	38.9	5	1	AL14_CARMA	P81817; carcinus ma
4	7	38.9	5	1	FARP_ARTTR	P41853; artiopposthi
5	7	38.9	5	1	PRCT_PERAM	P01373; periplaneta
6	7	38.9	5	1	PSK_DAUCA	P58261; daucus caro
7	5	27.8	4	1	FLRF_HIRME	P42551; hirudo medi
8	5	27.8	4	1	FMRF_MACNI	P01162; macrocallis
9	5	27.8	4	1	TUFT_HUMAN	P01858; homo sapien
10	5	27.8	5	1	BIOB_SALTY	P12678; salmonella
11	5	27.8	3	1	UF01_MOUSE	P38839; mus musculu
12	4	22.2	3	1	GRWM_HUMAN	P01157; homo sapien
13	3	16.7	4	1	ACH1_ACHFU	P35904; achatina fu
14	3	16.7	5	1	PAP2_PARMA	P81864; pardachirus
15	3	16.7	5	1	SUGA_ACHDO	P19991; acheta dome
16	3	16.7	5	1	TPIS_CANFA	P54714; canis famil
17	3	16.7	5	1	UC22_MAIZE	P80628; zea mays (m
18	2	11.1	3	1	LUXE_VIBFI	P24272; vibrio fisc
19	2	11.1	3	1	THYL_PIG	P01151; sus scrofa
20	2	11.1	4	1	DCML_PSECH	P19916; pseudomonas
21	2	11.1	4	1	DCMS_PSECH	P19918; pseudomonas
22	2	11.1	5	1	BIOA_CITFR	P13071; citrobacter
23	2	11.1	5	1	BIOA_SALTY	P12677; salmonella
24	2	11.1	5	1	BIOB_CITFR	P12997; citrobacter
25	2	11.1	5	1	BPP7_BOTIN	P30425; bothrops in
26	2	11.1	5	1	TRM3_ECOLI	P13973; escherichia
27	1	5.6	4	1	EOSI_HUMAN	P02731; homo sapien
28	1	5.6	4	1	RM01_YEAST	P36515; saccharomyc
29	0	0.0	5	1	UXA4_CHLTR	P38005; chlamydia t

ALIGNMENTS

RESULT 1
FAR3_HIRME
ID   FAR3_HIRME   STANDARD;   PRT;   4 AA.
AC   P42562;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE YLRF-AMIDE.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC   FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES   4   4   AMIDATION.
SQ   SEQUENCE   4 AA;   598 MW; 69D4073B30000000 CRC64;

Qy   4 Y 4
         |
Db   1 Y 1

Query Match   38.9%;   Score 7; DB 1; Length 4;
Best Local Similarity   100.0%; Pred. No. 1e+05;
Matches   1; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 2
FAR4_HIRME
ID   FAR4_HIRME   STANDARD;   PRT;   4 AA.
AC   P42563;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE YMRF-AMIDE.
OS   Hirudo medicinalis (Medicinal leech).
OC   Eukaryota; Metazoa; Annelida; Clitellata; Hirudinida; Hirudinea;
OC   Arynchobdellida; Hirudiniformes; Hirudinidae; Hirudo.
OX   NCBI_TaxID=6421;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=92195954; PubMed=1686933;
RA   Evans B.D., Pohl J., Kartsonis M.A., Calabrese R.L.;
RT   "Identification of RFamide neuropeptides in the medicinal leech.";
RL   Peptides 12:897-908(1991).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC   FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES   4   4   AMIDATION.
SQ   SEQUENCE   4 AA;   616 MW; 69D4068B30000000 CRC64;

Qy   4 Y 4
         |
Db   1 Y 1

Query Match   38.9%;   Score 7; DB 1; Length 4;
Best Local Similarity   100.0%; Pred. No. 1e+05;
Matches   1; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

RESULT 3
AL14_CARMA
ID   AL14_CARMA   STANDARD;   PRT;   5 AA.
AC   P81817;

```
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CARCINUSTATIN 14.
OS   Carcinus maenas (Common shore crab) (green crab).
OC   Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC   Eumalacostraca; Eucarida; Decapoda; Pleocyemata; Brachyura;
OC   Eubrachyura; Portunoidea; Portunidae; Carcinus.
OX   NCBI_TaxID=6759;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Cerebral ganglion, and Thoracic ganglion;
RX   MEDLINE=98121193; PubMed=9461295;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Jaros P.P.,
RA   Thorpe A.;
RT   "Isolation and identification of multiple neuropeptides of the
RT   allatostatin superfamily in the shore crab Carcinus maenas.";
RL   Eur. J. Biochem. 250:727-734(1997).
CC   -!- FUNCTION: MAY ACT AS A NEUROTRANSMITTER OR NEUROMODULATOR.
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation; Multigene family.
FT   MOD_RES       5      5       AMIDATION (POTENTIAL).
SQ   SEQUENCE   5 AA;   586 MW;  6728779D5AB300000 CRC64;

Query Match             38.9%;  Score 7;  DB 1;  Length 5;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 Y 4
              |
Db          1 Y 1


RESULT 4
FARP_ARTTR
ID   FARP_ARTTR     STANDARD;      PRT;   5 AA.
AC   P41853;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE RYIRF-AMIDE.
OS   Artioposthia triangulata.
OC   Eukaryota; Metazoa; Platyhelminthes; Turbellarian Platyhelminthes;
OC   Rhabditophora; Seriata; Tricladida; Terricola; Geoplanidae;
OC   Arthurdendyus.
OX   NCBI_TaxID=132421;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=94211927; PubMed=7909164;
RA   Maule A.G., Shaw C., Halton D.W., Curry W.J., Thim L.;
RT   "RYIRFamide: a turbellarian FMRFamide-related peptide (FaRP).";
RL   Regul. Pept. 50:37-43(1994).
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       5      5       AMIDATION.
SQ   SEQUENCE   5 AA;   754 MW;  69D4004B44600000 CRC64;

Query Match             38.9%;  Score 7;  DB 1;  Length 5;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 Y 4
              |
Db          2 Y 2


RESULT 5
PRCT_PERAM
ID   PRCT_PERAM     STANDARD;      PRT;   5 AA.
AC   P01373;
DT   21-JUL-1986 (Rel. 01, Created)
DT   30-JUL-1986 (Rel. 39, Last sequence update)
DT   01-FEB-1995 (Rel. 31, Last annotation update)
DE   PROCTOLIN.
OS   Periplaneta americana (American cockroach),
OS   Limulus polyphemus (Atlantic horseshoe crab), and
OS   Carcinus maenas (Common shore crab) (Green crab).
OC   Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Orthopteroidea; Dictyoptera; Blattaria;
OC   Blattoidea; Blattidae; Periplaneta.
OX   NCBI_TaxID=6978, 6850, 6759;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=P. americana;
RX   MEDLINE=76074708; PubMed=576;
RA   Starratt A.N., Brown B.E.;
RT   "Structure of the pentapeptide proctolin, a proposed neurotransmitter
RT   in insects.";
RL   Life Sci. 17:1253-1256(1975).
RN   [2]
RP   BIOLOGICAL SOURCE.
RC   SPECIES=P. americana;
RX   MEDLINE=81225665; PubMed=6113690;
RA   O'Shea M., Adams M.E.;
RT   "Pentapeptide (proctolin) associated with an identified neuron.";
RL   Science 213:567-569(1981).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=L. polyphemus;
RX   MEDLINE=90287800; PubMed=23561151;
RA   Groome J.R., Tillinghast E.K., Townley M.A., Vetrovs A.,
RA   Watson W.H. III, Hunt D.F., Griffin P.R., Alexander J.E.,
RA   Shabanowitz J.;
RT   "identification of proctolin in the central nervous system of the
RT   horseshoe crab, Limulus polyphemus.";
RL   Peptides 11:205-211(1990).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=c. maenas;
RX   MEDLINE=86232789; PubMed=2872661;
RA   Stangier J., Dircksen H., Keller R.;
RT   "identification and immunocytochemical localization of proctolin in
RT   periceardial organs of the shore crab, Carcinus maenas.";
RL   Peptides 7:67-72(1986).
CC   -!- FUNCTION: STIMULATES CARDIAC OUTPUT AND HINDGUT MOTILITY,
CC       MODULATES VISCERAL AND SKELETAL MUSCLE IN MANY ARTHROPODS.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE LATERAL WHITE NEURONS AND IN
CC       THE CRAB PERICARDIAL ORGANS.
DR   PIR; A01644; HGROHA.
DR   PIR; A01411; A60411.
KW   Neuropeptide.
SQ   SEQUENCE   5 AA;   649 MW;  71B76738A4600000 CRC64;

Query Match             38.9%;  Score 7;  DB 1;  Length 5;
Best Local Similarity   100.0%;  Pred. No. 1e+05;
Matches    1;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 Y 4
              |
Db          2 Y 2


RESULT 6
PSK_DAUCA
ID   PSK_DAUCA      STANDARD;      PRT;   5 AA.
AC   P58261;
DT   20-AUG-2001 (Rel. 40, Created)
DT   20-AUG-2001 (Rel. 40, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   PHYTOSULFOKINE-ALPHA (PSK-ALPHA) [CONTAINS: PHYTOSULFOKINE-BETA (PSK-
DE   BETA)].
OS   Daucus carota (Carrot).
```

```
              GenCore version 4.5
```

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:32:30 ; Search time 25.34 Seconds
(without alignments)
15.030 Million cell updates/sec

Title: us-09-281-760c-1
Perfect score: 18
Sequence: 1 LXXYR 5
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 205

Database : PIR_68:*
  1: pir1:*
  2: pir2:*
  3: pir3:*
  4: pir4:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

ninum DB seq length: 0
ximum DB seq length: 5

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	12	66.7	3	2	A22565	R-phycoerythrin al
2	12	66.7	5	2	F22565	R-phycoerythrin ga
3	8	44.4	4	2	I61883	protamine P1 - ora
4	8	44.4	4	2	I37013	protamine P1 - Cer
5	8	44.4	4	2	I84439	protamine P1 - sav
6	8	44.4	4	2	B37325	pap fimbrial regul
7	8	44.4	5	2	I39964	ribosomal protein
8	8	44.4	5	2	I39966	ribosomal protein
9	8	44.4	5	2	I39965	ribosomal protein
10	8	44.4	5	2	I40469	dnaZX-like protein
11	7	38.9	3	2	A43391	TRH-like tripeptid
12	7	38.9	3	2	S68328	blood cell protein
13	7	38.9	2	2	A32039	tyrosine-melanocyt
14	7	38.9	2	2	A37832	phenol 2-monooxyge
15	7	38.9	2	2	S09478	globulin IV alpha
16	7	38.9	2	2	PF0240	Ig heavy chain CRD
17	7	38.9	2	2	S43959	Ig mu chain V regi
18	7	38.9	5	1	HOROHA	proctolin - Americ
19	7	38.9	5	2	E60274	major protein anti
20	7	38.9	5	2	PQ0009	angiotensin-conver
21	7	38.9	5	2	PQ0689	photosystem I 10.4
22	7	38.9	5	2	B61445	Leu-enkephalin - b
23	7	38.9	5	2	A61445	Met-enkephalin - b
24	7	38.9	5	2	A60411	proctolin - Atlant
25	7	38.9	5	2	PT0278	Ig heavy chain CRD
26	7	38.9	5	2	S68326	blood cell protein
27	7	38.9	5	4	A58728	serrawettin W2 - S
28	5	27.8	3	2	PQ0010	angiotensin-conver
29	5	27.8	4	1	ECXAA	antho-RFamide neur
30	5	27.8	4	2	A02147	phagocytosis-stimu
31	5	27.8	4	2	ECNK	cardioexcitatory n
32	5	27.8	4	2	D41654	hypothetical prote
33	5	27.8	4	2	I40870	phospholipase C (E
34	5	27.8	4	2	JQ1273	neuropeptide Antho
35	5	27.8	4	2	A35779	neuropeptide Antho
36	5	27.8	4	2	A25844	antho-RF amide neu
37	5	27.8	4	2	A60418	FMRFamide - polych
38	5	27.8	4	2	PT0721	T-cell receptor be
39	5	27.8	4	2	S47552	ubiquitin - rat
40	5	27.8	4	2	JN0862	peptidyl-dipeptida
41	5	27.8	4	2	I40702	primase - Citrobac
42	5	27.8	4	2	A44955	alkanal monooxygen
43	5	27.8	4	2	D60274	major protein anti
44	5	27.8	4	2	B22565	R-phycoerythrin al
45	5	27.8	4	2	T14910	hypothetical prote
46	5	27.8	4	2	A60803	neuropeptide - sea
47	5	27.8	4	2	S53595	hypothetical prote
48	5	27.8	4	2	PT0295	seminal plasma pro
49	5	27.8	4	2	S62883	Ig heavy chain CRD
50	5	27.8	4	2	PT0513	T-cell receptor be

ALIGNMENTS

RESULT 1
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KLO>

Query Match 66.7%; Best Local Similarity 100.0%; Score 12; DB 2; Length 3;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YR 5
       ||
Db 2 YR 3

RESULT 2
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: F22565; MUID:85182601
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

Query Match 66.7%; Best Local Similarity 100.0%; Score 12; DB 2; Length 5;
Matches 2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 4 YR 5
       ||
Db 4 YR 5
```

RESULT 3
I61883
protamine P1 - orangutan (fragment)
C;Species: Pongo pygmaeus (orangutan)
C;Date: 06-Sep-1996
C;Accession: I61883
R;Qualit, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-end
A;Reference number: I37013; MUID:9404 0810
A;Accession: I61883
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: EMBL:Z12146; NID:g38156; PIDN:CAA78130.1; PID:g4379372

Query Match   44.4%; Score 8; DB 2; Length 4;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MARY 4

RESULT 4
I37013
protamine P1 - Cercopithecus patas (fragment)
C;Species: Cercopithecus patas
C;Date: 19-Mar-1997 #sequence_revision 07-Nov-1997 #text_change 21-Jul-2000
C;Accession: I37013
R;Qualit, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-end
A;Reference number: I37013; MUID:9404 0810
A;Accession: I37013
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: EMBL:Z12150; NID:g22814; PIDN:CAA78134.1; PID:g4377415

Query Match   44.4%; Score 8; DB 2; Length 4;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MARY 4

RESULT 5
I84439
protamine P1 - savannah baboon (fragment)
C;Species: Papio hamadryas doguera (savannah baboon)
C;Date: 19-Mar-1997 #sequence_revision 07-Nov-1997 #text_change 21-Jul-2000
C;Accession: I84439
R;Qualit, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-end
A;Reference number: I37013; MUID:9404 0810
A;Accession: I84439
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-4 <RES>
A;Cross-references: EMBL:Z12147; NID:g38134; PIDN:CAA78131.1; PID:g4379349

Query Match   44.4%; Score 8; DB 2; Length 4;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MARY 4

RESULT 6
B37325
pap fimbrial regulatory protein papI - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Date: 11-Sep-1992 #sequence_revision 11-Sep-1992 #text_change 23-Mar-1993
C;Accession: B37325
J. Braaten, B. A.; Blyn, L. B.; Skinner, B. S.; Low, D. A.
J. Bacteriol. 173, 1789-1800, 1991
A;Title: Evidence for a methylation-blocking factor (mbf) locus involved in pap pilus
A;Reference number: A37325; MUID:9115 4136
A;Accession: B37325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <BRA>
A;Cross-references: GB:M63747

Query Match   44.4%; Score 8; DB 2; Length 5;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MARY 4

RESULT 7
I39964
ribosomal protein S4 - Bacillus circulans (fragment)
C;Species: Bacillus circulans
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996
C;Accession: I39964
R;Grundy, F.J.; Henkin, T. M.
J. Bacteriol. 174, 6763-6770, 1992
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.
A;Reference number: I39963; MUID:93015735
A;Accession: I39964
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-5 <RES>
A;Cross-references: GB:M99041; NID:g143471
A;Genetics:
A;Gene: rpsD

Query Match   44.4%; Score 8; DB 2; Length 5;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MSEY 4

RESULT 8
I39966
ribosomal protein S4 - Bacillus licheniformis (fragment)
C;Species: Bacillus licheniformis
C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 19-Jul-1996
C;Accession: I39966
R;Grundy, F.J.; Henkin, T.M.
J. Bacteriol. 174, 6763-6770, 1992
A;Title: Characterization of the Bacillus subtilis rpsD regulatory target site.
A;Reference number: I39966; MUID:93015735
A;Accession: I39966
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA

Query Match   44.4%; Score 8; DB 2; Length 5;
Best Local Similarity   25.0%; Pred. No. 2.2e+05;
Matches   1; Conservative   1; Mismatches   2; Indels   0; Gaps   0;

Qy   1 LXXY 4
          : |
Db   1 MARY 4

us-09-281-760c-5.closed.rai

OM protein - protein search, using sw model

GenCore version 4.5

Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:51:47 ; Search time 12.9 seconds
                (without alignments)
                19.189 Million cell updates/sec

Title: US-09-281-760C-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5 .
Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 68711
Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	28	44.4	11	2	US-08-486-839-12	Sequence 12, Appl
2	28	44.4	11	3	US-09-151-011-11	Sequence 11, Appl
3	28	44.4	11	4	US-09-343-623-11	Sequence 11, Appl
4	26	41.3	8	1	US-08-014-426-16	Sequence 16, Appl
5	26	41.3	8	1	US-08-014-426-39	Sequence 39, Appl
6	26	41.3	8	5	PCT-US94-01319-16	Sequence 16, Appl
7	26	41.3	8	5	PCT-US94-01319-39	Sequence 39, Appl
8	25	39.7	4	4	US-09-082-593-11	Sequence 11, Appl
9	25	39.7	4	5	PCT-US94-01319-11	Sequence 11, Appl
10	25	39.7	8	1	US-08-014-426-17	Sequence 17, Appl
11	25	39.7	8	1	US-08-014-426-23	Sequence 23, Appl
12	25	39.7	8	1	US-08-014-426-38	Sequence 38, Appl
13	25	39.7	8	5	PCT-US94-01319-17	Sequence 17, Appl
14	25	39.7	8	5	PCT-US94-01319-23	Sequence 23, Appl
15	24	38.1	6	5	PCT-US94-01319-38	Sequence 38, Appl
16	24	38.1	6	1	US-08-482-528-127	Sequence 127, Appl
17	24	38.1	8	1	US-08-482-228-127	Sequence 127, Appl
18	24	38.1	8	1	US-08-014-426-29	Sequence 29, Appl
19	24	38.1	4	1	US-08-444-818-567	Sequence 567, Appl
20	24	38.1	5	1	US-08-444-818-566	Sequence 566, App
21	24	38.1	8	5	PCT-US94-01319-29	Sequence 29, Appl
22	24	38.1	9	2	US-08-725-736D-19	Sequence 19, Appl
23	24	38.1	9	3	US-09-162-368B-19	Sequence 19, Appl
24	24	38.1	9	4	US-09-161-877B-19	Sequence 19, Appl
25	24	38.1	9	4	US-09-171-705-65	Sequence 65, Appl
26	24	38.1	10	1	US-08-218-025A-122	Sequence 122, App
27	24	38.1	10	2	US-08-655-704B-2	Sequence 2, Appl
28	24	38.1	10	3	US-09-107-755-2	Sequence 2, Appli
29	24	38.1	11	2	US-08-486-839-11	Sequence 11, Appl
30	24	38.1	11	3	US-09-151-011-11	Sequence 11, Appl
31	24	38.1	11	4	US-09-343-623-11	Sequence 11, Appl
32	23	36.5	6	1	US-08-487-006-17	Sequence 17, Appl
33	23	36.5	6	2	US-08-209-261B-6	Sequence 6, Appli
34	23	36.5	6	2	US-08-317-310A-19	Sequence 19, Appl
35	23	36.5	6	2	US-08-488-659A-17	Sequence 17, Appl
36	23	36.5	7	5	PCT-US95-13041-19	Sequence 19, Appl
37	23	36.5	7	1	US-08-230-473-2	Sequence 2, Appl
38	23	36.5	8	1	US-08-014-426-21	Sequence 21, Appl
39	23	36.5	8	1	US-08-014-426-28	Sequence 28, Appl
40	23	36.5	8	1	US-08-014-426-36	Sequence 36, Appl
41	23	36.5	8	1	US-08-014-426-37	Sequence 37, Appl
42	23	36.5	8	1	US-08-403-634-11	Sequence 11, Appl
43	23	36.5	8	1	US-08-403-634-14	Sequence 14, Appl
44	23	36.5	8	1	US-08-189-331-162	Sequence 162, Appl
45	23	36.5	8	1	US-08-189-331-163	Sequence 162, App
46	23	36.5	8	2	US-08-199-331-163	Sequence 163, App
47	23	36.5	8	2	US-08-471-068-162	Sequence 162, App
48	23	36.5	8	2	US-08-471-068-163	Sequence 163, App
49	23	36.5	8	4	US-08-913-441B-11	Sequence 11, Appl
50	23	36.5	8	4	US-08-913-441B-14	Sequence 14, Appl
	23	36.5	8	4	US-09-026-904-7	Sequence 7, Appl

ALIGNMENTS

RESULT 1
US-08-486-839-12
Sequence 12, Application US/08486839
Patent No. 5928928
GENERAL INFORMATION:
    APPLICANT:
        TITLE OF INVENTION: A human chitinase, its recombinant
        TITLE OF INVENTION: production, its use for decomposing chitin, its use
        TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
    NUMBER OF SEQUENCES: 16
    CORRESPONDENCE ADDRESS:
        ADDRESSEE: Hoffmann & Baron
        STREET: 350 Jericho Turnpike
        CITY: Jericho
        STATE: New York
        COUNTRY: United States of America
        ZIP: 11758
    COMPUTER READABLE FORM:
        MEDIUM TYPE: Floppy disk
        COMPUTER: IBM PC compatible
        OPERATING SYSTEM: PC-DOS/MS-DOS
        SOFTWARE: Patentin Release #1.0, Version #1.30
    CURRENT APPLICATION DATA:
        APPLICATION NUMBER: US/08/486,839
        FILING DATE: 07 - June - 1995
    ATTORNEY/AGENT INFORMATION:
        NAME: Baron, Ronald J.
        REGISTRATION NUMBER: 29,281
        REFERENCE/DOCKET NUMBER: 294-26
    TELECOMMUNICATION INFORMATION:
        TELEPHONE: (516) 822-3550
        TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 12:
    SEQUENCE CHARACTERISTICS:
        LENGTH: 11 amino acids
        TYPE: amino acid
        STRANDEDNESS: unknown
        TOPOLOGY: unknown
        MOLECULE TYPE: protein
        HYPOTHETICAL: NO
US-08-486-839-12

Query Match   44.4%; Score 28; DB 2; Length 11;

RESULT 2
US-09-151-011-12
Sequence 12, Application US/09151011
Patent No. 6057142
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A Human Chitinase, Its Recombinant
TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron, LLP
STREET: 6900 Jericho Turnpike
CITY: Syosset
STATE: New York
COUNTRY: United States of America
ZIP: 11791
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/151,011
FILING DATE: 10 - September - 1998
ATTORNEY/AGENT INFORMATION:
NAME: Morliss, Robert C.
REGISTRATION NUMBER: 42,910
REFERENCE/DOCKET NUMBER: 294-32 DIV
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-151-011-12

Qy   1 GMNLTW 6
        |:||| |
Db   3 GlNLDW 8

Best Local Similarity   66.7%;   Pred. No. 34;
Matches   4;   Conservative   1;   Mismatches   1;   Indels   0;   Gaps   0;

STREET: 350 Jericho Turnpike
CITY: Jericho
STATE: New York
COUNTRY: United States of America
ZIP: 11758
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/343,623
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/486,839
FILING DATE: 07-June-1995
ATTORNEY/AGENT INFORMATION:
NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 11 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-343-623-12

Qy   1 GMNLTW 6
        |:||| |
Db   3 GlNLDW 8

Query Match   44.4%;   Score 28;   DB 4;   Length 11;
Best Local Similarity   66.7%;   Pred. No. 34;
Matches   4;   Conservative   1;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT 4
US-08-014-426-16
Sequence 16, Application US/08014426
Patent No. 5512435
GENERAL INFORMATION:
APPLICANT: Renschler, Markus F.
APPLICANT: Levy, Ronald
APPLICANT: Bhatt, Ramesh
APPLICANT: Dower, William
TITLE OF INVENTION: RECEPTOR-BINDING ANTIPROLIFERATIVE
TITLE OF INVENTION: PEPTIDES
NUMBER OF SEQUENCES: 65
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend Khourie and Crew
STREET: 379 Lytton Avenue
CITY: Palo Alto
STATE: California
COUNTRY: US
ZIP: 94301
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/014,426
FILING DATE: 05-FEB-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

RESULT 3
US-09-343-623-12
Sequence 12, Application US/0933623
Patent No. 6303118
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: A human chitinase, its recombinant
TITLE OF INVENTION: production, its use for decomposing chitin, its use
TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann & Baron

Query Match   44.4%;   Score 28;   DB 3;   Length 11;
Best Local Similarity   66.7%;   Pred. No. 34;
Matches   4;   Conservative   1;   Mismatches   1;   Indels   0;   Gaps   0;

Qy   1 GMNLTW 6
        |:||| |
Db   3 GlNLDW 8

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:55:15 ; Search time 25.77 Seconds
(without alignments)
62.437 Million cell updates/sec

Title: US-09-281-760c-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues
Total number of hits satisfying chosen parameters: 1277

Minimum DB seq length: 0
Maximum DB seq length: 11

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database : SPTREMBL_17:*
          1: sp_archea:*
          2: sp_bacteria:*
          3: sp_fungi:*
          4: sp_plant:*
          5: sp_invertebrate:*
          6: sp_mammal:*
          7: sp_mhc:*
          8: sp_organelle:*
          9: sp_phage:*
         10: sp_rodent:*
         11: sp_virus:*
         12: sp_human:*
         13: sp_vertebrate:*
         14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

ult No.	Score	% Query Match	Length	DB	ID	Description
1	25	39.7	9	4	Q9UMA0	Q9uma0 homo sapien
2	25	39.7	9	12	O85723	O85723 simian sarc
3	22	34.9	9	4	Q9UC36	Q9uc36 homo sapien
4	20	31.7	11	7	O77893	O77893 oreochromis
5	19	30.2	8	8	Q9VD02	Q9vd02 terranatos
6	18	28.6	8	2	Q9RC2	Q9rc2 borrelia bu
7	18	28.6	9	4	O16386	O16386 homo sapien
8	18	28.6	9	4	Q9UK44	Q9uk44 homo sapien
9	18	28.6	9	4	Q9BFF9	Q9bff9 homo sapien
10	18	28.6	10	8	Q15342	Q15342 homo sapien
11	18	28.6	10	8	Q9MQ05	Q9mq05 podospora c
12	18	28.6	10	13	P82080	P82080 limnodynast
13	18	28.6	11	13	Q90735	Q90735 gallus gall
14	17.5	27.8	10	8	P92771	P92771 xenosaurus
15	17	27.0	9	5	P8203	P8203 bombyx mori
16	17	27.0	10	4	Q9UJ48	Q9uj48 homo sapien
17	17	27.0	10	5	026693	026693 pisaster oc
18	17	27.0	10	12	Q08082	Q08082 chimpanzee
19	17	27.0	11	4	Q9UE69	Q9ue69 homo sapien
20	16.5	26.2	10	8	O79891	O79891 crotaphytus
21	16.5	26.2	10	8	O79894	O79894 gambelia wi
22	16.5	26.2	10	8	O79909	O79909 sauromalus
23	16.5	26.2	10	8	Q92YV3	Q92yv3 dipsosaurus
24	16.5	26.2	10	8	Q9T8G5	Q9t8g5 liolaemus o
25	16	25.4	8	5	Q9VRD2	Q9vrd2 drosophila
26	16	25.4	8	11	Q9QVD3	Q9qvd3 rattus sp.
27	16	25.4	9	7	Q9R635	Q9r635 chlamydia t
28	16	25.4	9	2	Q9MW43	Q9mw43 homo sapien
29	16	25.4	9	12	O82622	O82622 avian infec
30	16	25.4	10	4	O54217	O54217 staphylococ
31	16	25.4	10	10	Q9UCU6	Q9ucu6 homo sapien
32	16	25.4	11	10	P82443	P82443 nicotiana t
33	16	25.4	11	8	Q9UEX7	Q9uex7 cnemidophor
34	15.5	24.6	10	8	P92616	P92616 cnemidophor
35	15.5	24.6	10	8	Q9TG98	Q9tg98 shinisaurus
36	15	23.8	10	7	P82065	P82065 litoria rub
37	15	23.8	8	4	Q9UMH9	Q9umh9 homo sapien
38	15	23.8	8	11	Q9UJ50	Q9uj50 homo sapien
39	15	23.8	8	6	062721	062721 rattus norv
40	15	23.8	9	6	Q9TU33	Q9tu33 canis famil
41	15	23.8	10	8	P92576	P92576 bipes bipor
42	15	23.8	10	11	P92632	P92632 eremias gra
43	15	23.8	10	11	Q9R0Z0	Q9r0z0 mus musculu
44	15	23.8	10	2	Q9PRU9	Q9pru9 sparus aura
45	15	23.8	11	2	047345	047345 escherichia
46	15	23.8	11	7	O77892	O77892 oreochromis
47	15	23.8	11	7	O77899	O77899 oreochromis
48	15	23.8	11	7	O77900	O77900 oreochromis
49	15	23.8	11	7	O77901	O77901 oreochromis
50	15	23.8	11	7	O77902	O77902 oreochromis

ALIGNMENTS

RESULT 1

Q9UMA0

ID   Q9UMA0   PRELIMINARY;   PRT;   9 AA.
AC   Q9UMA0;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   KIT PROTEIN (FRAGMENT).
GN   KIT.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94061059; PubMed=7694728;
RA   Spritz R.A., Holmes S.A., Berg S.Z., Nordlund J.J., Fukai K.;
RT   "A recurrent deletion in the KIT (mast/stem cell growth factor
RT   receptor) proto-oncogene is a frequent cause of human piebaldism.";
RL   Hum. Mol. Genet. 2:1499-1500(1993).
DR   EMBL; S67686; AAD13996.1; -.
NON_TER   1   1
SQ   SEQUENCE   9 AA;   1182 MW; 0BC504032361B5AB CRC64;

Query Match   39.7%; Score 25; DB 4; Length 9;
Best Local Similarity   66.7%;
Matches   4; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   5 TWYRES 10
            || ||:
Db   4 TWKREN 9

RESULT 2

O85723

ID   O85723   PRELIMINARY;   PRT;   9 AA.

```
AC   Q85723;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   SIS ONCOGENE (FRAGMENT).
OS   Simian sarcoma virus.
OC   Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX   NCBI_TaxID=11817;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84106822; PubMed=6319011;
RA   Devare S.G., Shatzman A.R., Robbins K.C., Rosenberg M., Aaronson S.A.;
RT   "Expression of the PDGF-related transforming protein of simian sarcoma
RT   virus in E. coli.";
RL   Cell 36:43-49(1984).
DR   EMBL; K01473; AAA46816.1; -.
FT   NON_TER      9      9
SQ   SEQUENCE   9 AA;  1048 MW;  9C53A866C361A731 CRC64;

Qy   2  MNLTW  6
        |||||
Db   1  MTLTW  5

Query Match             39.7%;  Score 25;  DB 12;  Length 9;
Best Local Similarity   80.0%;  Pred. No. 4.7e+05;
Matches    4;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

RESULT 3
ID   Q9UC36
AC   Q9UC36;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   28 KDA HEAT SHOCK PROTEIN HOMOLOG FRAGMENT 1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
SEQUENCE.
RX   MEDLINE=92218434; PubMed=1560006;
RA   Kato K., Shinohara H., Goto S., Inaguma Y., Morishita R., Asano T.;
RT   "Copurification of small heat shock protein with alpha B crystallin
RT   from human skeletal muscle.";
RL   J. Biol. Chem. 267:7718-7725(1992).
SQ   SEQUENCE   9 AA;  1220 MW;  2693315B1F77B43 CRC64;

Qy   6  WYRESK  11
        |  | ||
Db   2  WYPHSR  7

Query Match             34.9%;  Score 22;  DB 4;  Length 9;
Best Local Similarity   50.0%;  Pred. No. 4.7e+05;
Matches    3;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

RESULT 4
ID   077893
AC   077893;
DT   01-NOV-1998 (TrEMBLrel. 08, Created)
DT   01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   MHC CLASS II B LOCUS 10 (FRAGMENT).
OS   Oreochromis niloticus (Nile tilapia) (Tilapia nilotica).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Labroidei;
OC   Cichlidae; Oreochromis.
```

```
OX   NCBI_TaxID=8128;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=98315113; PubMed=9649539;
RA   Malaga-Trillo E., McAndrew B., Vincek V., Zaleska-Rutczynska Z.,
RA   Sueltmann H., Figueroa F., Klein J.;
RT   "Linkage relationships and haplotype polymorphism among cichlid Mhc
RT   class II B loci.";
RL   Genetics 149:1527-1547(1998).
DR   EMBL; AF050003; AAC41342.1; -.
KW   MHC.
FT   NON_TER      1      1
FT   NON_TER     11     11
SQ   SEQUENCE   11 AA;  1296 MW;  68775B73786B572B CRC64;

Qy   6  WIRE  9
        ||||
Db   1  WIRE  4

Query Match             31.7%;  Score 20;  DB 7;  Length 11;
Best Local Similarity   75.0%;  Pred. No. 2.5e+03;
Matches    3;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

RESULT 5
ID   Q9TD02
AC   Q9TD02;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   CYTOCHROME C OXIDASE SUBUNIT I (FRAGMENT).
OS   Terranatos dolichopterus.
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC   Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC   Cyprinodontiformes; Aplocheilidae; Rivulinae; Terranatos.
OG   Mitochondrion.
OX   NCBI_TaxID=81836;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Hrbek T., Larson A.;
RT   "The evolution of diapause in the killifish family Rivulidae
RT   (Atherinomorpha, Cyprinodontiformes): A molecular phylogenetic and
RT   biogeographic perspective.";
RL   Evolution 53:1200-1216(1999).
DR   EMBL; AF092421; AAF03041.1; -.
OG   Mitochondrion.
FT   NON_TER      8      8
SQ   SEQUENCE   8 AA;  1084 MW;  F0C9D3640DD44056 CRC64;

Qy   2  MNLTW   7
        |  |:
Db   2  INARWF  7

Query Match             30.2%;  Score 19;  DB 8;  Length 8;
Best Local Similarity   33.3%;  Pred. No. 4.7e+05;
Matches    2;  Conservative    2;  Mismatches    2;  Indels    0;  Gaps    0;

RESULT 6
ID   Q9R9C2
AC   Q9R9C2;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PLASMID CP32-3, POSSIBLE PARTITION PROTEINS, COMPLETE CDS (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-3.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
```

us-09-281-760c-5.closed.rsp

```
                                  GenCore version 4.5
                      Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:           April 21, 2002, 10:56:21 ; Search time 10.33 Seconds
                                              (without alignments)
                                              39.043 Million cell updates/sec

Title:            US-09-281-760C-5
Perfect score:    63
Sequence:         1 GMNLTWYRESK 11

Scoring table:
Scoring table:    BLOSUM62
                  Gapop 10.0 , Gapext 0.5

Searched:         100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters:       388

imum DB seq length: 0
imum DB seq length: 11

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :        SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

```
                          SUMMARIES

        %   Query
Result  Score  Match Length DB  ID         Description
NO.
-----   -----  ----------- --  --         -----------
 1      26    41.3    10  1  APE_CAPGI       capnocytoph
 2      19    30.2    10  1  UPA9_HUMAN      homo sapien
 3      18    28.6    10  1  GON2_CHICK      gallus gall
 4      17    27.0     8  1  UPAA_HUMAN      homo sapien
 5      16    25.4     9  1  HUTU_KLEAE      klebsiella
 6      16    25.4     9  1  LMIP_LOCMI      locusta mig
 7      15    23.8     9  1  FARP_CALSI      callinectes
 8      15    23.8     9  1  NEF_HV1Z8       human immun
 9      15    23.8    10  1  GON1_PETMA      petromyzon
10      15    23.8    10  1  P04378 PNEU_RAT  rattus norv
11      14    22.2    10  1  RL16_ACHLA      acholeplasm
12      14    22.2    10  1  URE3_MORMO      morganella
13      14    22.2    10  1  UXA6_CHLTR      chlamydia t
14      13.5  21.4     8  1  LCK3_LEUMA      leucophaea
15      13    20.6     8  1  ACI_THUAL       thunnus alb
16      13    20.6     8  1  PLP_BRANA       brassica na
17      13    20.6    10  1  BPP2_BOTJA      bothrops ja
18      13    20.6    10  1  GON3_PETMA      petromyzon
19      13    20.6    10  1  LABA_JATMU      jatropha mu
20      13    20.6    11  1  TOP1_SALTY      salmonella
21      13    20.6    11  1  LPW_THETH       thermus aqu
22      13    20.6    11  1  TKN2_UPERU      uperoleia r
23      12    19.0     6  1  LOK1_LOCMI      locusta mig
24      12    19.0     7  1  IGAO_DACDE      dactylium d
25      12    19.0     7  1  WWA1_ACHFU      achatina fu
26      12    19.0     7  1  WWA2_ACHFU      achatina fu
27      12    19.0     8  1  AKH_LIBAU       libellula a
28      12    19.0     8  1  LCK1_LEUMA      leucophaea
29      12    19.0     8  1  LCK2_LEUMA      leucophaea
30      12    19.0     8  1  LCK4_LEUMA      leucophaea
31      12    19.0     8  1  LCK5_LEUMA      leucophaea
32      12    19.0     8  1  LCK6_LEUMA      leucophaea
33      12    19.0     8  1  LCK7_LEUMA      leucophaea
34      12    19.0    12  1  LCK8_LEUMA      leucophaea
35      12    19.0    12  1  UHA2_HUMAN      homo sapien
36      12    19.0    10  1  AKHX_LOCMI      locusta mig
37      12    19.0    10  1  COXO_RAT        rattus norv
38      12    19.0    10  1  HTFI_ROMMI      romalea mic
39      12    19.0    10  1  NO40_TOBAC      nicotiana t
40      12    19.0    10  1  SYK_CAMJP       campylobact
41      12    19.0    10  1  TKUI_UREUN      urechis uni
42      12    19.0    10  1  CEP1_ACHFU      achatina fu
43      12    19.0    11  1  ULAG_HUMAN      homo sapien
44      11    17.5     5  1  BPP7_BOTIN      bothrops in
45      11    17.5     5  1  UF01_MOUSE      mus musculu
46      11    17.5     7  1  WWA3_ACHFU      achatina fu
47      11    17.5     8  1  AKHG_GRYBI      gryllus bim
48      11    17.5     8  1  MELML          melolontha
49      11    17.5     8  1  AKH_TABAT       tabanus atr
50      11    17.5     8  1  ALL5_CALVO      calliphora
```

```
                          ALIGNMENTS

RESULT 1
APE_CAPGI
ID   APE_CAPGI        STANDARD;      PRT;      10 AA.
AC   P80474;
DT   01-OCT-1996 (Rel. 34, Created)
DT   01-OCT-1996 (Rel. 34, Last sequence update)
DT   01-OCT-1996 (Rel. 34, Last annotation update)
DE   AMINOPEPTIDASE (EC 3.4.11.-) (FRAGMENT).
OS   Capnocytophaga gingivalis.
OC   Bacteria; CFB group; Flavobacteria; Flavobacteriaceae; Capnocytophaga.
OX   NCBI_TaxID=1017;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=ATCC 33624;
RX   MEDLINE=96118234; PubMed=8574402;
RA   Spratt D.A., Greenman J., Schaffer A.G.;
RT   "Capnocytophaga gingivalis aminopeptidase: a potential virulence
RT   factor.";
RL   Microbiology 141:3087-3093(1995).
CC   -!- FUNCTION: AMINOPEPTIDASE WHICH HYDROLYZES SUBSTRATES WITH FREE N-
CC       TERMINAL AMINO ACID RESIDUES BUT NOT N-TERMINAL BLOCKED ONES.
CC   -!- NUTRITION AND PATHOGENESIS OF THE ORGANISM IN THE HUMAN ORAL
CC       CAVITY.
CC   -!- COFACTOR: REQUIRES MAGNESIUM OR CALCIUM.
KW   Hydrolase; Aminopeptidase; Magnesium; Calcium.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1306 MW;  00C0A6DB43772694 CRC64;
     GMNLTWY 7

Query Match             41.3%;  Score 26;  DB 1;  Length 10;
Best Local Similarity   50.0%;  Pred. No. 54;
Matches    3;  Conservative    2;  Mismatches    1;  Indels   0;  Gaps   0;

Qy          2 MNLTWY 7
              :||: ||
Db          2 VNMLWY 7

RESULT 2
UPA9_HUMAN
ID   UPA9_HUMAN       STANDARD;      PRT;      10 AA.
AC   P30095;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 35) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

```
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1450097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 7.2, ITS MW IS: 15 KDA.
DR   SWISS-2DPAGE; P30095; HUMAN.
FT   NON_TER       1      1
FT   NON_TER      10     10
SQ   SEQUENCE   10 AA;  1233 MW;  37AD72B409C581B7 CRC64;

Qy       3 NLTWYRESK 11
Db       2 NPTMFRDNE 10

Query Match             30.2%;  Score 19;  DB 1;  Length 10;
  Best Local Similarity   33.3%;  Pred. No. 9e+02;
  Matches    3;  Conservative    4;  Mismatches    2;  Indels    0;  Gaps    0;


RESULT 3
GON2_CHICK
ID   GON2_CHICK     STANDARD;      PRT;    10 AA.
AC   P37043; P20408; P81750;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   GONADOLIBERIN II (GONADOTROPIN-RELEASING HORMONE II) (GNRH-II)
DE   (LH-RH II) (LUIBERIN II).
OS   Gallus gallus (Chicken).
OS   Alligator mississippiensis (American alligator),
OS   Squalus acanthias (Spiny dogfish),
OS   Hydrolagus colliei (Spotted ratfish) (Pacific ratfish), and
OS   Clupea pallasii (Pacific herring).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031, 8496, 7797, 7873, 30724;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=Chicken; TISSUE=Hypothalamus;
RX   MEDLINE=84222059; PubMed=6427779;
RA   Miyamoto K., Hasegawa Y., Nomura M., Igarashi M., Kangawa K.,
RA   Matsuo H.;
RT   "Identification of the second gonadotropin-releasing hormone in
RT   chicken hypothalamus: evidence that gonadotropin secretion is
RT   probably controlled by two distinct gonadotropin-releasing hormones
RT   in avian species.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:3874-3878(1984).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=A.mississippiensis; TISSUE=Brain;
RX   MEDLINE=91352338; PubMed=1882082;
RA   Lovejoy D.A., Fischer W.H., Parker D.B., McCrory J.E., Park M.,
RA   Lance V., Swanson P., Rivier J.E., Sherwood N.M.;
RT   "Primary structure of two forms of gonadotropin-releasing hormone
RT   from brains of the American alligator (Alligator mississippiensis).";
RL   Regul. Pept. 33:105-116(1991).
RN   [3]
RP   SEQUENCE.
RC   SPECIES=S.acanthias; TISSUE=Brain;
RX   MEDLINE=92335300; PubMed=1631133;
RA   Lovejoy D.A., Fischer W.H., Ngamvongchon S., Craig A.G.,
RA   Nahorniak C.S., Peter R.E., Rivier J.E., Sherwood N.M.;
RT   "Distinct sequence of gonadotropin-releasing hormone (GnRH) in
RT   dogfish brain provides insight into GnRH evolution.";
RL   Proc. Natl. Acad. Sci. U.S.A. 89:6373-6377(1992).
RN   [4]
RP   SEQUENCE.
RC   SPECIES=H.colliei; TISSUE=Brain;
RX   MEDLINE=91340067; PubMed=1678723;
RA   Lovejoy D.A., Sherwood N.M., Fischer W.H., Jackson B.C., Rivier J.E.,
RA   Lee T.;
RT   "Primary structure of gonadotropin-releasing hormone from the brain
RT   of a holocephalan (ratfish: Hydrolagus colliei).";
RL   Gen. Comp. Endocrinol. 82:152-161(1991).
RN   [5]
RP   SEQUENCE, AND FUNCTION.
RC   SPECIES=C.pallasii; TISSUE=Brain, and Pituitary;
RX   MEDLINE=20114351; PubMed=10650929;
RA   Carolsfeld J., Powell J.F.F., Park M., Fischer W.H., Craig A.G.,
RA   Chang J.P., Rivier J.E., Sherwood N.M.;
RT   "Primary structure and function of three gonadotropin-releasing
RT   hormones, including a novel form, from an ancient teleost, herring.";
RL   Endocrinology 141:505-512(2000).
CC   -!- FUNCTION: STIMULATES THE SECRETION OF GONADOTROPINS.
CC   -!- SIMILARITY: BELONGS TO THE GNRH FAMILY.
DR   PIR; B60066; RHAQ2.
DR   PIR; A61126; A61126.
DR   PIR; B46030; B46030.
DR   PROSITE; PS00473; GNRH; 1.
DR   Pfam; PF00446; GnRH; 1.
DR   InterPro; IPR002012; GnRH.
KW   Hormone; Amidation; Hypothalamus.
FT   MOD_RES       1      1       PYROLIDONE CARBOXYLIC ACID.
FT   MOD_RES      10     10       AMIDATION.
SQ   SEQUENCE   10 AA;  1254 MW;  284B2E437871F5A3 CRC64;

Qy       6 WY 7
Db       7 WY 8

Query Match             28.6%;  Score 18;  DB 1;  Length 10;
  Best Local Similarity  100.0%;  Pred. No. 1.3e+03;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;


RESULT 4
UPAA_HUMAN
ID   UPAA_HUMAN     STANDARD;      PRT;     8 AA.
AC   P30096;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 36) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;
RX   MEDLINE=93092937; PubMed=1450097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 7, ITS MW IS: 12 KDA.
DR   SWISS-2DPAGE; P30096; HUMAN.
FT   NON_TER       1      1
FT   NON_TER       5      5
FT   VARIANT       5      5       F -> P. /FTId=VAR_000004.
SQ   SEQUENCE    8 AA;   909 MW;  86677B59D1A72042 CRC64;
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:54:10 ; Search time 19.63 Seconds
                                          (without alignments)
                                          42.686 Million cell updates/sec

Title: US-09-281-760C-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 1324

Database :   PIR_68:*
                1: pir1:*
                2: pir2:*
                3: pir3:*
                4: pir4:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

nimum DB seq length: 0
ximum DB seq length: 11

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	Match	DB	ID	Description
1	25	39.7	7	7	2	PN0649	alpha-dextrin endo
2	23	36.5	11	11	2	A33571	follistatin - bovi
3	21	33.3	10	10	2	F49033	T-cell receptor ga
4	19	30.2	10	7	2	C54226	light-harvesting p
5	18	28.6	7	7	2	S21230	dermorphin (Trp-4,
6	18	28.6	9	9	2	A61230	calsequestrin, car
7	18	28.6	9	2	2	S78426	52.5K protein - sp
8	18	28.6	10	1	2	RHAQ2	gonadoliberin II -
9	18	28.6	10	1	1	A61126	gonadoliberin II -
10	18	28.6	10	1	2	B46030	gonadoliberin - sp
11	18	28.6	10	2	2	PR0291	Ig heavy chain CRD
12	18	28.6	11	2	2	PT0310	Ig heavy chain CRD
13	18	28.6	10	2	2	PT0302	Ig heavy chain CRD
14	17.5	27.8	10	2	2	T14219	cytochrome-c oxida
15	17	27.0	10	2	2	F06589	sperm-activating p
16	17	27.0	10	2	2	B38887	T-cell receptor ga
17	16.5	26.2	10	2	2	T17057	cytochrome-c oxida
18	16.5	26.2	10	2	2	T12303	cytochrome-c oxida
19	16.5	26.2	10	2	2	T17060	cytochrome-c oxida
20	16.5	26.2	10	2	2	T17072	cytochrome-c oxida
21	16.5	26.2	9	1	2	AKLQIM	locustamyoinhibiti
22	16	25.4	9	9	2	C36730	hutU protein - Kle
23	16	25.4	9	2	2	PT0272	Ig heavy chain CRD
24	16	25.4	10	2	2	H60787	sperm-activating p
25	16	25.4	10	2	2	D60788	sperm-activating p
26	16	25.4	10	2	2	B60787	sperm-activating p
27	16	25.4	10	2	2	A60788	sperm-activating p
28	16	25.4	10	2	2	G60527	sperm-activating p
29	16	25.4	10	2	2	C60589	sperm-activating p
30	16	25.4	10	2	2	D60588	sperm-activating p
31	16	25.4	11	2	2	C53652	rhlR protein - Pse
32	16	25.4	11	2	2	JQ2307	hypothetical 1.5K
33	16	25.4	11	2	2	JQ2317	hypothetical 1.5K
34	16	25.4	11	2	2	A29806	acidic proline-ric
35	15.5	25.4	11	2	2	A49037	TcR gamma V-J regi
36	15	24.6	10	2	2	T13976	cytochrome-c oxida
37	15	23.8	10	2	2	F22565	R-phycoerythrin ga
38	15	23.8	7	2	2	A60803	neuropeptide - sea
39	15	23.8	5	2	1	PX0008	glucuronosyltransf
40	15	23.8	9	2	2	RHLMGS	phosphoenolpyruvat
41	15	23.8	10	2	2	A47593	mercury resistance
42	15	23.8	10	2	2	PC2172	gonadoliberin - se
43	15	23.8	10	2	2	G60589	triacylglycerol li
44	15	23.8	10	2	2	A60589	sperm-activating p
45	15	23.8	10	2	2	B39308	sperm-activating p
46	15	23.8	10	2	2	PT0289	glycine reductase
47	15	23.8	10	2	2	PH0923	Ig heavy chain CRD
48	15	23.8	10	2	2	T13838	T-cell receptor be
49	15	23.8	10	2	2	T14019	cytochrome-c oxida
50	15	23.8	10	2	2		cytochrome-c oxida

ALIGNMENTS

RESULT   1
PN0649
alpha-dextrin endo-1,6-alpha-glucosidase (EC 3.2.1.41) - Bacillus sp. (strain S-1) (f
C;Species: Bacillus sp.
C;Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 07-May-1999
C;Accession: PN0649
R;Kim, C.H.; Choi, H.I.; Lee, D.S.
R;Kim, C.H.; Choi, H.I.; Lee, D.S.
Biosci. Biotechnol. Biochem. 57, 1632-1637, 1993
A;Title: Purification and biochemical properties of an alkaline pullulanase from alka
A;Reference number: PN0649; MUID:94080025
A;Accession: PN0649
A;Molecule type: protein
A;Residues: 1-7 <KIM>
C;Comment: This enzyme is used together with glucoamylase to improve the efficiency of
ntent in high maltose syrups.
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match   39.7%;   Score 25;   DB 2;   Length 7;
Best Local Similarity   33.3%;   Pred. No. 2.2e+05;
Matches   2;   Conservative.   4;   Mismatches 0;   Indels 0;   Gaps 0;

Qy   2 MNLTWY 7
          :|:|:|:
Db   2 LNMSWF 7

RESULT   2
A33571
follistatin - bovine (fragment)
C;Species: Bos primigenius taurus (cattle)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 30-Sep-1993
C;Accession: A33571
R;Gospodarowicz, D.; Lau, K.
Biochem. Biophys. Res. Commun. 165, 292-298, 1989
A;Title: Pituitary follicular cells secrete both vascular endothelial growth factor a
A;Reference number: A33571; MUID:90073725
A;Accession: A33571
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-11 <GOS>

Query Match   36.5%;   Score 23;   DB 2;   Length 11;
Best Local Similarity   50.0%;   Pred. No. 4.6e+02;
Matches   3;   Conservative.   2;   Mismatches 1;   Indels 0;   Gaps 0;

RESULT 3

```
Qy     6 WYRESK 11
          || |::|
Db     4 WLRQAK 9
```

F49033
T-cell receptor gamma chain V-D-J region - human (fragment)
C;Species: Homo sapiens (man)
C;Date: 19-Dec-1993 #sequence_revision 17-Mar-2000 #text_change 17-Mar-2000
C;Accession: F49033
R;Morita, C.T.; Verma, S.; Aparicio, P.; Martinez, C.; Spits, H.; Brenner, M.B.
Eur. J. Immunol. 21, 2999-3007, 1991
A;Title: Functionally distinct subsets of human gamma/delta T cells.
A;Reference number: A49033; MUID:92083926
A;Accession: F49033
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-10 <MOR>
A;Cross-references: GB:S72605; NID:g240700; PIDN:AAB20632.1; PID:g240701
A;Note: sequence extracted from NCBI backbone (NCBIN:72605, NCBIP:72606)
C;Keywords: T-cell receptor

```
Query Match     33.3%;  Score 21;  DB 2;  Length 10;
Best Local Similarity   50.0%;  Pred. No. 9.2e+02;
Matches    2;  Conservative    2;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     6 WYRE 9
          || :::
Db     7 WYKK 10
```

RESULT 4

C54426
light-harvesting protein B-830 alpha-3 chain - Chromatium purpuratum (fragment)
C;Species: Chromatium purpuratum
C;Date: 05-Jan-1996 #sequence_revision 05-Jan-1996
C;Accession: C54426
R;Kerfeld, C.A.; Yeates, T.O.; Thornber, J.P.
Biochemistry 33, 2178-2184, 1994
A;Title: Purification and characterization of the peripheral antenna of the purple-sulfu
A;Reference number: A54226; MUID:94162224
A;Accession: C54426
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-10 <KER>
A;Keywords: antenna complex; light-harvesting polypeptide

```
Query Match     30.2%;  Score 19;  DB 2;  Length 10;
Best Local Similarity  100.0%;  Pred. No. 2e+03;
Matches    4;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     8 RESK 11
          |||
Db     1 RESK 4
```

RESULT 5

S21230
dermorphin (Trp-4, Asn-7) [validated] - two-colored leaf frog (fragment)
C;Species: Phyllomedusa bicolor (two-colored leaf frog)
C;Date: 19-Mar-1997 #sequence_revision 10-Oct-1997 #text_change 18-Aug-2000
C;Accession: S21230
R;Mignogna, G.; Severini, C.; Simmaco, M.; Negri, L.; Falconieri Erspamer, G.; Krell, G.
FEBS Lett. 302, 151-154, 1992
A;Title: Identification and characterization of two dermorphins from skin extracts of th
A;Reference number: S21152; MUID:92339502
A;Accession: S21230
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-7 <MIG>
C;Superfamily: dermorphin precursor; dermorphin precursor amino-terminal homology

```
Query Match     28.6%;  Score 18;  DB 2;  Length 7;
Best Local Similarity  100.0%;  Pred. No. 2.2e+05;
Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy     6 WY 7
          ||
Db     4 WY 5
```

RESULT 6

A61230
calsequestrin, cardiac and slow skeletal muscle - northern leopard frog (fragment)
N;Alternate names: 58K dihydropyridine-binding protein; aspartactin; calmitine; lamin
C;Species: Rana pipiens (northern leopard frog)
C;Date: 03-May-1994 #sequence_revision 03-May-1994 #text_change 16-Jul-1999
C;Accession: A61230
R;McLeod, A.G.; Shen, A.C.Y.; Campbell, K.P.; Michalak, M.; Jorgensen, A.O.
Circ. Res. 69, 344-359, 1991
A;Title: Frog cardiac calsequestrin. Identification, characterization, and subcellula
rdium.
A;Reference number: A61230; MUID:91316784
A;Accession: A61230
A;Molecule type: protein
A;Residues: 1-9 <MCL>
C;Comment: Calsequestrin is a high-capacity and moderate-affinity calcium binding pro
C;Comment: Calsequestrin acts as a calcium buffer, and the release of calcium bound t
C;Comment: The fast skeletal muscle isoform of calsequestrin can be phosphorylated in
C;Superfamily: calsequestrin
C;Keywords: calcium binding; cardiac muscle; glycoprotein; heart; phosphoprotein; ske

```
Query Match     28.6%;  Score 18;  DB 2;  Length 9;
Best Local Similarity   42.9%;  Pred. No. 2.2e+05;
Matches    3;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy     1 GMNLTWY 7
          |:| |
Db     3 GINFPTY 9
```

RESULT 7

S78426
52.5K protein - spiny lobster (fragment)
C;Species: Panulirus argus (spiny lobster)
C;Date: 14-Apr-1998 #sequence_revision 24-Apr-1998 #text_change 19-May-2000
C;Accession: S78426
R;James, M.O.; Boyle, S.M.; Trapido-Rosenthal, H.G.; Smith, W.C.; Greenberg, R.M.; Sh
Arch. Biochem. Biophys. 329, 31-38, 1996
A;Title: cDNA and protein sequence of a major form of P450, CYP2L, in the hepatopancr
A;Reference number: S68856; MUID:96201120
A;Accession: S78426
A;Molecule type: protein
A;Residues: 1-9 <JAM>
A;Experimental source: hepatopancreas microsomes

```
Query Match     28.6%;  Score 18;  DB 2;  Length 9;
Best Local Similarity   40.0%;  Pred. No. 2.2e+05;
Matches    2;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy     5 TWYRE 9
          || :
Db     1 TWIKK 5
```

RESULT 8

RHAO2
gonadoliberin II - American alligator
gonadoliberin II - American alligator
N;Alternate names: gonadotropin-releasing hormone II
C;Species: Alligator mississippiensis (American alligator)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:48:46 ; Search time 12.88 Seconds
                   (without alignments)
                   12.230 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7

Scoring table: BLOSUM62
                   Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 35098

Post-processing: Minimum Match 0%
                   Maximum Match 100%
                   Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 7

Database :   Issued_Patents_AA:*
                   1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
                   2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
                   3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
                   4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
                   5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
                   6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	62.9	7	4	US-09-258-754-22	Sequence 22, Appli
2	22	62.9	7	1	US-09-042-107-22	Sequence 22, Appl
3	20	57.1	7	1	US-08-105-416-36	Sequence 36, Appl
4	20	57.1	7	2	US-08-473-656A-36	Sequence 36, Appl
5	18	51.4	5	4	US-09-020-880-78	Sequence 78, Appl
6	18	51.4	7	2	US-08-437-116A-11	Sequence 11, Appl
7	18	51.4	7	2	US-08-530-792D-11	Sequence 11, Appl
8	17	48.6	3	2	US-08-968-676-27	Sequence 27, Appl
9	17	48.6	4	2	US-08-968-676-26	Sequence 26, Appl
10	17	48.6	5	4	US-08-968-676-25	Sequence 25, Appl
11	17	48.6	5	5	US-09-020-880-77	Sequence 77, Appl
12	17	48.6	5	1	US-08-727-045A-1	Sequence 1, Appl
13	17	48.6	6	2	US-08-166-930-1	Sequence 1, Appl
14	17	48.6	6	2	US-08-968-676-24	Sequence 24, Appl
15	17	48.6	6	1	US-08-727-045A-1	Sequence 1, Appl
16	17	48.6	6	6	572855-18 Patent No. 5472855	Sequence 12, Appl
17	17	48.6	1	6	US-08-335-198-12	Sequence 12, Appl
18	17	48.6	7	6	US-08-968-676-22	Sequence 22, Appl
19	17	48.6	7	4	US-09-258-754-271	Sequence 271, App
20	17	48.6	7	4	US-09-042-107-271	Sequence 271, App
21	16	45.7	7	5	US-08-789-333F-77	Sequence 77, Appl
22	16	45.7	6	2	US-08-519-103-4	Sequence 4, Appl
23	16	45.7	6	2	US-08-204-417A-4	Sequence 4, Appl
24	16	45.7	6	2	US-08-177-109A-8	Sequence 8, Appl
25	16	45.7	6	2	US-08-177-109A-38	Sequence 38, Appl
26	16	45.7	6	2	US-08-687-706-8	Sequence 8, Appl
27	16	45.7	6	2	US-08-687-706-38	Sequence 38, Appl
28	16	45.7	6	4	US-09-018-635-4	Sequence 4, Appli
29	16	45.7	7	1	US-09-401-512-74	Sequence 74, Appl
30	16	45.7	7	1	US-08-401-512-76	Sequence 76, Appl
31	16	45.7	7	1	US-08-264-093-25	Sequence 25, Appl
32	16	45.7	7	1	US-08-346-333-84	Sequence 84, Appl
33	16	45.7	7	1	US-08-290-301-81	Sequence 81, Appl
34	16	45.7	7	3	US-09-188-579-105	Sequence 105, Appl
35	16	45.7	7	1	US-09-273-565-60	Sequence 60, Appl
36	16	45.7	7	4	US-09-226-012-109	Sequence 109, App
37	16	45.7	7	4	US-09-226-012-111	Sequence 111, App
38	16	45.7	7	1	US-09-475-316A-43	Sequence 43, Appl
39	16	45.7	7	5	US-09-315-444-105	Sequence 105, Appl
40	16	45.7	7	5	PCT-US91-07506-84	Sequence 84, Appl
41	16	42.9	5	1	US-09-313-075A-22	Sequence 22, Appl
42	15	42.9	5	1	US-08-460-343B-47	Sequence 47, Appl
43	15	42.9	5	2	US-08-398-028B-47	Sequence 47, Appl
44	15	42.9	5	1	US-08-504-265B-47	Sequence 47, Appl
45	15	42.9	5	3	US-09-130-663-7	Sequence 7, Appli
46	15	42.9	5	2	US-09-649-100-19	Sequence 19, Appl
47	15	42.9	5	3	US-08-649-100-35	Sequence 35, Appl
48	15	42.9	5	4	US-09-020-880-80	Sequence 80, Appl
49	15	42.9	5	4	US-09-203-623-22	Sequence 22, Appl
50	15	42.9	5	4	US-09-432-335-7	Sequence 7, Appli

ALIGNMENTS

RESULT 1
US-09-258-754-22
Sequence 22, Application US/09258754
Patent No. 6174687
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki
APPLICANT: Pasqualini, Renata
APPLICANT: Rajotte, Daniel
TITLE OF INVENTION: Methods of Identifying Lung Homing Molecules Using
TITLE OF INVENTION: Membrane Dipeptidase
FILE REFERENCE: P-LJ 3443
CURRENT APPLICATION NUMBER: US/09/258,754
CURRENT FILING DATE: 1999-02-26
EARLIER APPLICATION NUMBER: 09/042,107
EARLIER FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 452
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
LENGTH: 7
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic

Query Match   62.9%; Score 22; DB 4; Length 7;
Best Local Similarity 66.7%; Pred. No. 1.6e+05;
Matches   4; Conservative   1; Mismatches   1; Indels   0; Gaps   0;

Qy   1 TLLEYR 6
           : ||||
Db   2 SFLEYR 7

RESULT 2
US-09-042-107-22
Sequence 22, Application US/09042107
Patent No. 6232287
GENERAL INFORMATION:
APPLICANT: Ruoslahti, Erkki
APPLICANT: Pasqualini, Renata
TITLE OF INVENTION: Molecules that Home to Various Selected Organs or
TITLE OF INVENTION: Tissues
FILE REFERENCE: P-LJ 2892

CURRENT APPLICATION NUMBER: US/09/042,107
CURRENT FILING DATE: 1998-03-13
NUMBER OF SEQ ID NOS: 436
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 22
  LENGTH: 7
  TYPE: PRT
  ORGANISM: Artificial Sequence
  FEATURE:
  OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-042-107-22

Qy   1 TLLEYR 6
        : |||
Db   2 SFLEYR 7

Query Match   62.9%;   Score 22;   DB 4;   Length 7;
Best Local Similarity   66.7%;   Pred. No. 1.6e+05;
Matches   4;   Conservative   1;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT 3
US-08-105-416-36
Sequence 36, Application US/08105416
Patent No. 5639958
GENERAL INFORMATION:
  APPLICANT: Olsson, Lennart
  APPLICANT: Goodenow, Robert S
  APPLICANT: Goldstein, Avram
  TITLE OF INVENTION: Class I MHC Modulation of Surface
  TITLE OF INVENTION: Receptor Activity
  NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
  STREET: Four Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: CA
  COUNTRY: US
  ZIP: 94111
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/105,416
  FILING DATE: 12-AUG-1993
  CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
  NAME: Rowland, Bertram I.
  REGISTRATION NUMBER: 20015
  REFERENCE/DOCKET NUMBER: A55115-4/BIR
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-781-1989
  TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 36:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 7 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: peptide
US-08-105-416-36

Qy   1 TLLEY 5
        ||| |
Db   3 TLLRY 7

Query Match   57.1%;   Score 20;   DB 1;   Length 7;
Best Local Similarity   80.0%;   Pred. No. 1.6e+05;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT 4
US-08-473-656A-36
Sequence 36, Application US/08473656A
Patent No. 5853999
GENERAL INFORMATION:
  APPLICANT: Olsson, Lennart
  APPLICANT: Goodenow, Robert S
  APPLICANT: Goldstein, Avram
  TITLE OF INVENTION: Class I MHC Modulation of Surface
  TITLE OF INVENTION: Receptor Activity
  NUMBER OF SEQUENCES: 40
CORRESPONDENCE ADDRESS:
  ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert
  STREET: Four Embarcadero Center, Suite 3400
  CITY: San Francisco
  STATE: CA
  COUNTRY: US
  ZIP: 94111
COMPUTER READABLE FORM:
  MEDIUM TYPE: Floppy disk
  COMPUTER: IBM PC compatible
  OPERATING SYSTEM: PC-DOS/MS-DOS
  SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/473,656A
  FILING DATE:
PRIOR APPLICATION DATA:
  CLASSIFICATION: 435
  APPLICATION NUMBER: US 08/105,416
  FILING DATE: 12-AUG-1993
ATTORNEY/AGENT INFORMATION:
  NAME: Rowland, Bertram I.
  REGISTRATION NUMBER: 20015
  REFERENCE/DOCKET NUMBER: A55115-4/BIR
TELECOMMUNICATION INFORMATION:
  TELEPHONE: 415-781-1989
  TELEFAX: 415-398-3249
INFORMATION FOR SEQ ID NO: 36:
  SEQUENCE CHARACTERISTICS:
  LENGTH: 7 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
  MOLECULE TYPE: peptide
US-08-473-656A-36

Qy   1 TLLEY 5
        ||| |
Db   3 TLLRY 7

Query Match   57.1%;   Score 20;   DB 2;   Length 7;
Best Local Similarity   80.0%;   Pred. No. 1.6e+05;
Matches   4;   Conservative   0;   Mismatches   1;   Indels   0;   Gaps   0;

RESULT 5
US-09-020-880-78
Sequence 78, Application US/09020880A
Patent No. 6136558
GENERAL INFORMATION:
  APPLICANT: Genentech, Inc.
  APPLICANT: Ballinger, Marcus D.
  APPLICANT: Jones, Jennifer T.
  APPLICANT: Fairbrother, Wayne J.
  APPLICANT: Silkowski, Mark X.
  APPLICANT: Wells, James A.
  TITLE OF INVENTION: HEREGULIN VARIANTS
  FILE REFERENCE: 14918-7200CON1
  CURRENT APPLICATION NUMBER: US/09/020,880A
  CURRENT FILING DATE: 1998-02-09

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:50:47 ; Search time 25.72 Seconds
(without alignments)
39.810 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5
Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 65

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 50 summaries

Database :
SPTREMBL_17:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

imum DB seq length: 0
imum DB seq length: 7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

lt NO.	Score	% Query Match	Length	DB	ID	Description
1	37.1	13	7	12	Q9YQ10	Q9YQ10 porcine tra
2	28.6	10	5	13	P82071	P82071 litoria rub
3	28.6	10	5	13	P82072	P82072 litoria rub
4	28.6	10	7	2	Q54248	Q54248 streptomyce
5	22.9	8	6	4	Q08720	Q08720 homo sapien
6	22.9	8	7	8	O98866	O98866 spinacia ol
7	22.9	8	7	8	O99182	O99182 gnatholebia
8	22.9	8	12		O66205	O66205 porcine tra
9	20.0	7	2		P72081	P72081 nocardia la
10	20.0	7	2		O50556	O50556 actinobacil
11	20.0	7	4		Q15897	Q15897 homo sapien
12	20.0	7	5		Q28742	Q28742 oryctolagus
13	20.0	7	6		Q94421	Q94421 psathyrosta
14	20.0	7	8		P92385	P92385 hordeum mar
15	20.0	7	8		P92210	P92210 agropyron c
16	20.0	7	8		P92214	P92214 amblyopyrum
17	20.0	7	8		P92218	P92218 australopyr
18	20.0	7	8		P92221	P92221 bromus iner
19	20.0	7	8		P92226	P92226 crithopsis
20	20.0	7	8		P92372	P92372 haynaldia v
21	20.0	7	8		P92381	P92381 hordeum bra
22	20.0	7	8		P92387	P92387 hordeum bra
23	20.0	7	8		P92390	P92390 henardia p
24	20.0	7	8		P92393	P92393 heteranthel
25	20.0	7	8		P92425	P92425 hordeum vul
26	20.0	7	8		P92427	P92427 pseudoroegn
27	20.0	7	8		P92430	P92430 peridictyon
28	20.0	7	8		P92442	P92442 aegilops ta
29	20.0	7	8		P92440	P92440 taeniatheru
30	20.0	7	8		P92403	P92403 thinopyrum
31	20.0	7	12		Q9YVE3	Q9YVE3 lophopyrum
32	20.0	7	12		Q9YIR0	Q9YIR0 human adeno
33	20.0	7	12		Q9YIQ9	Q9YIQ9 human adeno
34	20.0	7	13		P82065	P82065 litoria rub
35	17.1	7	13		P82101	P82101 litoria rub
36	17.1	7	11		Q08433	Q08433 rattus norv
37	17.1	6	13		P82070	P82070 litoria rub
38	17.1	6	13		O34028	O34028 pseudomonas
39	17.1	2	2		P70804	P70804 azotobacter
40	17.1	7	2		Q95945	Q95945 saccharomyc
41	17.1	7	8		P82445	P82445 nicotiana t
42	17.1	7	12		Q66113	Q66113 cherry leaf
43	17.1	7	13		Q99007	Q99007 hordeum vul
44	14.3	5	13		Q99007	Q99007 litoria rub
45	14.3	5	13		P82073	P82073 litoria rub
46	14.3	5	13		P82099	P82099 litoria rub
47	14.3	5	10		P82181	P82181 spinacia ol
48	14.3	5	10		P82182	P82182 spinacia ol
49	14.3	6	10		P82541	P82541 spinacia ol
50	14.3	6	13		P82096	P82096 litoria rub

ALIGNMENTS

RESULT 1
Q9YQ10

ID Q9YQ10 PRELIMINARY; PRT; 7 AA.
AC Q9YQ10;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-MAY-1999 (TrEMBLrel. 10, Last annotation update)
DE HYPOTHETICAL FUSION PROTEIN.
OS porcine transmissible gastroenteritis virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Nidovirales;
OC Coronaviridae; Coronavirus.
OX NCBI_TaxID=11149;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99099045; PubMed=9882359;
RA Izeta A., Smerdou C., Alonso S., Penzes Z., Mendez A., Plana-Duran J.,
RA Enjuanes L.;
RT "Replication and packaging of transmissible gastroenteritis
RT coronavirus-derived synthetic minigenomes.";
RL J. Virol. 73:1535-1545(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95159435; PubMed=7856095;
RA Eleouet J., Rasschaert D., Lambert P., Levy L., Vende P., Laude H.;
RT "Complete sequence (20 kilobases) of the polyprotein-encoding gene 1
RT of transmissible gastroenteritis virus.";
RL Virology 206:817-822(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88078100; PubMed=2825819;
RA Rasschaert D., Gelfi J., Laude H.;
RT "Enteric coronavirus TGEV: partial sequence of the genomic RNA, its
RT organization and expression.";
RL Biochimie 69:591-600(1987).
DR EMBL; AJ011482; CAA09625.1; -.
SQ SEQUENCE 7 AA; 927 MW; 69D6D7273B5726F0 CRC64;

```
Query Match             37.1%;  Score 13;  DB 12;  Length 7;
Best Local Similarity   50.0%;  Pred. No. 4.7e+05;
Matches   2;  Conservative   2;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    2  LLEY  5
         |||.:
Db    4  LLQF  7


RESULT   2
ID   P82071                    PRELIMINARY;   PRT;   5 AA.
AC   P82071;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 2.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella', the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=626; METHOD=FAB.
KW   Amphibian skin.
SQ   SEQUENCE   5 AA;   626 MW;  6DD9C9CB10300000 CRC64;


Query Match             33.3%;  Score 10;  DB 13;  Length 5;
Best Local Similarity   33.3%;  Pred. No. 4.7e+05;
Matches   2;  Conservative   2;  Mismatches   0;  Indels   0;  Gaps   0;

Qy    3  LEY  5
         |||
Db    1  IEF  3


RESULT   3
ID   P82072                    PRELIMINARY;   PRT;   5 AA.
AC   P82072;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   RUBELLIDIN 3.1.
OS   Litoria rubella (Desert tree frog).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Amphibia; Batrachia; Anura; Neobatrachia; Bufonoidea; Hylidae;
OC   Litoria.
OX   NCBI_TaxID=104895;
RN   [1]
RP   SEQUENCE, AND MASS SPECTROMETRY.
RC   TISSUE=SKIN SECRETION;
RA   Steinborner S.T., Wabnitz P.A., Waugh R.J., Bowie J.H., Gao C.,
RA   Tyler M.J., Wallace J.C.;
RT   "The structure of new peptides from the Australian red tree frog
RT   'Litoria rubella', the skin peptide profile as a probe for the study
RT   of evolutionary trends of amphibians.";
RL   Aust. J. Chem. 49:955-963(1996).
CC   -!- FUNCTION: CAERIDINS SHOW NEITHER NEUROPEPTIDE ACTIVITY NOR
CC       ANTIBIOTIC ACTIVITY.
```

```
CC   -!- TISSUE SPECIFICITY: SECRETED BY THE SKIN DORSAL GLANDS.
CC   -!- MASS SPECTROMETRY: MW=655; METHOD=FAB.
KW   Amphibian skin; Amidation.
MOD_RES  5   AMIDATION.
FT
SQ   SEQUENCE   5 AA;   656 MW;  71A9C9CB10300000 CRC64;

Qy    3  LLEY  5
         |||:
Db    1  IEF  3


Query Match             28.6%;  Score 10;  DB 13;  Length 5;
Best Local Similarity   33.3%;  Pred. No. 4.7e+05;
Matches   1;  Conservative   2;  Mismatches   0;  Indels   0;  Gaps   0;


RESULT   4
ID   Q54248                    PRELIMINARY;   PRT;   7 AA.
AC   Q54248;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1996 (TrEMBLrel. 01, Last annotation update)
DE   SECY & ADK GENES (FRAGMENT).
GN   RPLO.
OS   Streptomyces griseus.
OC   Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC   Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX   NCBI_TaxID=1911;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=N2-3-11;
RA   Poehling S., Piepersberg W., Wehmeier U.F.;
RL   Submitted (FEB-1996) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; X95915; CAA65160.1; -.
FT   NON_TER       1
SQ   SEQUENCE   7 AA;   760 MW;  72C72B01B2D1B2A0 CRC64;

Qy    1  TLLE  4
         ||:|
Db    2  TVTE  5


Query Match             28.6%;  Score 10;  DB 2;  Length 7;
Best Local Similarity   50.0%;  Pred. No. 4.7e+05;
Matches   1;  Conservative   1;  Mismatches   1;  Indels   0;  Gaps   0;


RESULT   5
ID   Q08720                    PRELIMINARY;   PRT;   6 AA.
AC   Q08720;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-NOV-1998 (TrEMBLrel. 08, Last annotation update)
DE   Y PROTEIN (FRAGMENT).
GN   CREB.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=93010691; PubMed=1396344;
RA   Waeber G., Habener J.F.;
RT   "Novel testis germ cell-specific transcript of the CREB gene contains
RT   an alternatively spliced exon with multiple in-frame stop codons.";
RL   Endocrinology 131:2010-2015(1992).
DR   EMBL; X68994; CAA48780.1; -.
FT   NON_TER       1
FT   NON_TER       6
SQ   SEQUENCE   6 AA;   695 MW;  6727?EB9C735D000 CRC64;
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:51:12 ; Search time 10.46 Seconds
                                            (without alignments)
                                            24.537 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7

Database :   SwissProt_39:*

Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 74

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	34.3	12	6	1	ASP2_LACSN	P82655 lactobacill
2	34.3	12	6	1	FAR2_ASCSU	P31890 ascaris suu
3	28.6	10	6	1	OVM_LEPDE	P42985 leptinotars
4	28.6	10	7	1	CCF1_ENTFA	P20104 enterococcu
5	25.7	9	7	1	HY7_PIG	P01153 sus scrofa
6	22.9	8	6	1	FARP_MONEX	P41966 moniezia ex
7	22.9	8	7	1	FAR1_ASCSU	P31889 ascaris suu
8	22.9	8	7	1	FAR5_HIRME	P42564 hirudo medi
9	22.9	8	7	1	GFRP_MOUSE	P99025 mus muscu
10	22.9	8	7	1	MSCL_SALTY	P39446 salmonella
11	20.0	7	4	1	FAR3_HIRME	P42562 hirudo medi
12	20.0	7	4	1	FAR4_HIRME	P42563 hirudo medi
13	20.0	7	4	1	FLRF_HIRME	P42561 hirudo medi
14	20.0	7	5	1	AL14_CARMA	P81817 carcinus ma
15	20.0	7	5	1	FARP_ARTTR	P41853 artiopostni
16	20.0	7	5	1	PRCT_PERAM	P01373 periplaneta
17	20.0	7	5	1	PSK_DAUCA	P58261 daucus caro
18	20.0	7	6	1	ACPH_RABIT	P25154 oryctolagus
19	20.0	7	6	1	ALL2_CARMA	P81805 carcinus ma
20	20.0	7	6	1	ALL3_CARMA	P81806 carcinus ma
21	20.0	7	6	1	ALL4_CARMA	P81807 carcinus ma
22	20.0	7	6	1	ALL5_CARMA	P81808 carcinus ma
23	20.0	7	6	1	ALL7_CYDPO	P82158 cydia pomon
24	20.0	7	6	1	CARP_MYTED	P10470 mytilus edu
25	20.0	7	6	1	FAR1_HELTI	P41871 helisoma tr
26	20.0	7	6	1	FAR1_PROCL	P38499 procambarus
27	20.0	7	6	1	FAR2_PROCL	P38498 procambarus
28	20.0	7	6	1	FAR3_HAECO	P81298 haemonchus
29	20.0	7	6	1	FAR3_PANRE	P41874 panagrellus
30	20.0	7	6	1	MNP1_LEPDE	P42984 leptinotars
31	20.0	7	6	1	MYOM_APLCA	P15513 aplysia cal
32	20.0	7	6	1	UN06_PINPS	P81675 pinus pinas
33	20.0	7	1		WWA1_ACHFU	P35919 achatina fu
34	17.1	6	6	1	TMOF_SARBU	P41495 sarcophaga
35	17.1	6	6	1	CIA_ENTFA	P11932 enterococcu
36	17.1	6	6	1	LANC_CARUI	P36960 carnobacter
37	17.1	6	6	1	UF03_MOUSE	P38641 mus musculu
38	17.1	6	7	1	WWA2_ACHFU	P35920 achatina fu
39	14.3	5	4	1	DCML_PSECH	P19916 pseudomonas
40	14.3	5	4	1	DCMS_PSECH	P19918 pseudomonas
41	14.3	5	4	1	EOSI_HUMAN	P02731 homo sapien
42	14.3	5	4	1	FMRF_MACNI	P01162 macrocallis
43	14.3	5	4	1	RM01_YEAST	P35515 saccharomyc
44	14.3	5	4	1	TUFT_HUMAN	P01858 homo sapien
45	14.3	5	4	1	BIOA_CITFR	P13071 citrobacter
46	14.3	5	5	1	BIOA_SALTY	P12677 salmonella
47	14.3	5	5	1	BIOB_CITFR	P12997 citrobacter
48	14.3	5	5	1	BIOB_SALTY	P12678 salmonella
49	14.3	5	5	1	TRM3_ECOLI	P13973 escherichia
50	14.3	5	1		UC22_MAIZE	P80628 zea mays (m

ALIGNMENTS

RESULT 1
ASP2_LACSN

ID   ASP2_LACSN   STANDARD;   PRT;   6 AA.
AC   P82655;
DT   20-AUG-2001 (Rel. 40, Created);
DT   20-AUG-2001 (Rel. 40, Last sequence update);
DT   20-AUG-2001 (Rel. 40, Last annotation update);
DE   ACID SHOCK PROTEIN 2 (FRAGMENT).
OS   Lactobacillus sanfranciscensis (Lactobacillus sanfrancisco).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;
OC   Lactobacillus.
OX   NCBI_TaxID=1625;
RN   [1]
RP   SEQUENCE.
RC   STRAIN=CB1;
RX   PubMed=11429463;
RA   De Angelis M., Bini L., Pallini V., Cocconcelli P.S., Gobbetti M.;
RT   "The acid-stress response in Lactobacillus sanfranciscensis CB1.";
RL   Microbiology 147:1863-1873(2001).
CC   -!- INDUCTION: OVEREXPRESSED IN ACID ENVIRONMENTS.
FT   NON_TER   6
SQ   SEQUENCE   6 AA;   778 MW;   6AA45B5B132A8000 CRC64;

Query Match   34.3%; Score 12; DB 1; Length 6;
Best Local Similarity 100.0%;
Matches   2; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy   4 EY 5
          ||
Db   3 EY 4

RESULT 2
FAR2_ASCSU

ID   FAR2_ASCSU   STANDARD;   PRT;   7 AA.
AC   P31890;
DT   01-JUL-1993 (Rel. 26, Created)
DT   01-JUL-1993 (Rel. 26, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   FMRFAMIDE-LIKE NEUROPEPTIDE AF2.
OS   Ascaris suum (Pig roundworm) (Ascaris lumbricoides), and
OS   Panagrellus redivivus.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC   Ascarididae; Ascaris.
OX   NCBI_TaxID=6253, 6233;
RN   [1]
RP   SEQUENCE.
RC   SPECIES=A.suum;
RX   MEDLINE=93324431; PubMed=8332542;
RA   Cowden C., Stretton A.O.W.;

```
RT   "AF2, an Ascaris neuropeptide: isolation, sequence, and bioactivity.";
RL   Peptides 14:423-430(1993).
RN   [2]
RP   SEQUENCE.
RC   SPECIES=P.redivivus;
RX   MEDLINE=95060098; PubMed=7970891;
RA   Maule A.G., Shaw C., Bowman J.W.;
RT   "The FMRFamide-like neuropeptide AF2 (Ascaris suum) is present in the
RT   free-living nematode, Panagrellus redivivus (Nematoda, Rhabditida).";
RL   Parasitology 109:351-356(1994).
CC   -!- FUNCTION: HAS EFFECTS ON MUSCLE TENSION.
CC   -!- TISSUE SPECIFICITY: FOUND IN THE NERVE CORDS AND A VARIETY OF
CC       GANGLIA PARTICULARLY IN THE ANTERIOR REGIONS.
CC   -!- SIMILARITY: BELONGS TO THE FARP (FMRFAMIDE RELATED PEPTIDE)
CC       FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES        7      7       AMIDATION.
SQ   SEQUENCE   7 AA;   992 MW;  69D4073B5B11E350 CRC64;

Query Match             34.3%;  Score 12;  DB 1;  Length 7;
Best Local Similarity  100.0%;  Pred. No. 1e+05;
Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 EY 5
              ||
Db          3 EY 4

RESULT 3
OVM_LEPDE
ID   OVM_LEPDE     STANDARD;      PRT;    6 AA.
AC   P42985;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   OVIDUCTAL MOTILITY STIMULATING PEPTIDE (LED-OVM).
OS   Leptinotarsa decemlineata (Colorado potato beetle).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;
OC   Chrysomelinae; Leptinotarsa.
OX   NCBI_TaxID=7539;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RX   MEDLINE=91271080; PubMed=2052497;
RA   Spittaels K., Schoofs L., Grauwels L., Smet H., van Damme J.,
RA   Proost P., Torrekens S., de Loof A.;
RT   "Isolation, identification and synthesis of novel oviductal motility
RT   stimulating head peptide in the Colorado potato beetle, Leptinotarsa
RT   decemlineata.";
RL   Peptides 12:31-36(1991).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
CC       OVIDUCT.
KW   Neuropeptide; Amidation.
FT   MOD_RES        6      6       AMIDATION.
SQ   SEQUENCE   6 AA;   720 MW;  6B07632B5DD03000 CRC64;

Query Match             28.6%;  Score 10;  DB 1;  Length 6;
Best Local Similarity   25.0%;  Pred. No. 1e+05;
Matches    1;  Conservative    2;  Mismatches    1;  Indels    0;  Gaps    0;

Qy          3 LEYR 6
              : |:
Db          1 IAYK 4

RESULT 4
CCF1_ENTFA
ID   CCF1_ENTFA    STANDARD;      PRT;    7 AA.
AC   P20104;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   01-FEB-1991 (Rel. 17, Last annotation update)
DE   SEX PHEROMONE CCF10.
OS   Enterococcus faecalis (Streptococcus faecalis).
OC   Bacteria; Firmicutes; Bacillus/Clostridium group; Enterococcaceae;
OC   Enterococcus.
OX   NCBI_TaxID=1351;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=89008313; PubMed=3139658;
RA   Mori M., Sakagami Y., Ishii Y., Isogai A., Kitada C., Fujino M.,
RA   Adsit J.C., Dunny G.M., Suzuki A.;
RT   "Structure of cCF10, a peptide sex pheromone which induces
RT   conjugative transfer of the Streptococcus faecalis tetracycline
RT   resistance plasmid, pCF10.";
RL   J. Biol. Chem. 263:14574-14578(1988).
CC   -!- FUNCTION: CCF10 IS INVOLVED IN THE CONJUGATIVE TRANSFER OF THE
CC       HEMOLYSIN PLASMID PCF10.
DR   PIR; A30812; A30812.
KW   Pheromone.
SQ   SEQUENCE   7 AA;   790 MW;  72C9D22C731B2C740 CRC64;

Query Match             28.6%;  Score 10;  DB 1;  Length 7;
Best Local Similarity   66.7%;  Pred. No. 1e+05;
Matches    2;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TLL 3
              ||
Db          3 TLV 5

RESULT 5
HY7_PIG
ID   HY7_PIG       STANDARD;      PRT;    7 AA.
AC   P01153;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   21-JUL-1986 (Rel. 01, Last annotation update)
DE   HYPOTHALAMIC HEPTAPEPTIDE.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=81213980; PubMed=6263778;
RA   Chang R.C.C., Huang W-Y., Arimura A., Redding T.W., Coy D.H.,
RA   Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT   "Isolation, structure and synthesis of a heptapeptide with in vitro
RT   ACTH-releasing activity from porcine hypothalamus.";
RL   Horm. Metab. Res. 13:228-232(1981).
DR   PIR; A01417; NYPG7.
SQ   SEQUENCE   7 AA;   957 MW;  632B45B1FB5059A0 CRC64;

Query Match             25.7%;  Score 9;  DB 1;  Length 7;
Best Local Similarity   50.0%;  Pred. No. 1e+05;
Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          5 YR 6
              |:
Db          6 YK 7

RESULT 6
FARP_MONEX
ID   FARP_MONEX    STANDARD;      PRT;    6 AA.
AC   P41966;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1997 (Rel. 35, Last annotation update)
```

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:49:07 ; Search time 19.61 Seconds
                 (without alignments)
                 27.191 Million cell updates/sec

Title: US-09-281-760C-4
Perfect score: 35
Sequence: 1 TLLEYRM 7
Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5
Searched: 219241 seqs, 76174552 residues
Total number of hits satisfying chosen parameters: 455

Database :   PIR_68:*
                 1: pir1:*
                 2: pir2:*
                 3: pir3:*
                 4: pir4:*

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 7

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	%Query Match	Length	DB	ID	Description
1	16	45.7	7	2	S25266	pilE protein - Esc
2	13	37.1	5	2	S69237	surface protein te
3	13	37.1	5	2	I37027	protamine P1 - gor
4	13	37.1	6	2	S71867	glutathione transf
5	12	34.3	3	2	A22565	R-phycoerythrin al
6	12	34.3	5	2	B37325	pap fimbrial regul
7	12	34.3	5	2	F22565	R-phycoerythrin ga
8	12	34.3	6	2	A41946	T-cell receptor ga
9	12	34.3	7	4	A38081	amine oxidase (cop
10	12	34.3	7	4	I56695	hypothetical protein
11	11	31.4	6	2	T11779	phosphoglycerate t
12	11	31.4	7	2	T71870	glutathione transf
13	11	31.4	7	2	B34818	vicilin 57K chain
14	10	28.6	5	2	I40469	dnaZX-like protein
15	10	28.6	6	2	S11556	hydrogensulfite re
16	10	28.6	7	2	A30812	sex pheromone cCF1
17	9	25.7	3	2	A43391	hydrogensulfite re
18	9	25.7	4	2	T13892	TRH-like tripeptid
19	9	25.7	4	2	E44823	cytochrome-c oxida
20	9	25.7	5	2	E42364	synaptosomal-assoc
21	9	25.7	5	2	S68326	flagellar protein
22	9	25.7	5	2	S11024	blood cell protein
23	9	25.7	6	2	B44510	hydrogensulfite re
24	9	25.7	6	2	A43766	hypothetical prote
25	9	25.7	6	2	JH0784	28K ubiquitin-immu
26	9	25.7	6	2	I65546	neuropeptide TE-6
27	9	25.7	7	1	NYPG7	MHC H2-L antigen -
28	9	25.7	7	2	JN0859	hypothalamic hepta
29	9	25.7	7	2	S78024	peptidyl-dipeptida
30	8	22.9	4	2	A37832	phenol 2-monooxyge
31	8	22.9	4	2	I61883	protamine P1 - ora
32	8	22.9	4	2	I37013	protamine P1 - Cer
33	8	22.9	4	2	I84439	protamine P1 - sav
34	8	22.9	5	2	I39964	ribosomal protein
35	8	22.9	5	2	I39966	ribosomal protein
36	8	22.9	5	2	I39965	ribosomal protein
37	8	22.9	5	2	PQ0689	photosystem I 10.4
38	8	22.9	5	2	A44692	fulicin - giant Af
39	8	22.9	5	2	A60986	collagen alpha 1(I
40	8	22.9	6	2	B56979	Y protein - human
41	8	22.9	6	2	I37263	N-formyl oligopept
42	8	22.9	6	2	I49808	glycoprotein compo
43	8	22.9	6	2	H43394	D-SP2.5 region - m
44	8	22.9	6	2	B33932	Ig mu chain D regi
45	8	22.9	6	2	I49424	cytotoxic T-lympho
46	8	22.9	6	2	I49421	laminin B1 - weste
47	8	22.9	6	2	A43129	neuropeptide GNFFR
48	8	22.9	6	2	S20446	elastase - Pseudom
49	8	22.9	7	2	S17976	glucose isomerase
50	8	22.9	7	2	A34026	acetylcholinestera

ALIGNMENTS

RESULT 1
S25266
pilE protein - Escherichia coli (fragment)
C;Species: Escherichia coli
C;Accession: S25266 #sequence_revision 28-May-1993 #text_change 11-Jun-1993
R;Dupuy, B.; Taha, M.K.; Possot, O.; Marchal, C.; Pugsley, A.P.
Mol. Microbiol. 6, 1887-1894, 1992
A;Title: PulO, a component of the pullulanase secretion pathway of Klebsiella oxytoca
A;Reference number: S25266; MUID:92374839
A;Accession: S25266
A;Molecule type: protein
A;Residues: 1-7 <DUP>
C;Genetics:
A;Gene: pilE

Query Match   45.7%;   Score 16;   DB 2;   Length 7;
Best Local Similarity   75.0%;   Pred. No. 2.2e+05;
Matches   3;   Conservative   1;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy    1  TLLE  4
         ||:|
Db    2  TLIE  5
```

RESULT 2
S69237
surface protein tetrabrachion heavy chain - Staphylothermus marinus (fragment)
C;Species: Staphylothermus marinus
C;Date: 04-Dec-1997 #sequence_revision 04-Dec-1997 #text_change 17-Mar-1999
C;Accession: S69237
R;Peters, J.; Nitsch, M.; Kuehlmorgen, B.; Golbik, R.; Lupas, A.; Kellermann, J.; Eng
J. Mol. Biol. 245, 385-401, 1995
A;Title: Tetrabrachion: a filamentous archaebacterial surface protein assembly of unu
A;Reference number: S69237; MUID:95139068
A;Accession: S69237
A;Molecule type: protein
A;Residues: 1-5 <PET>
A;Experimental source: strain F1, DSM 3639
A;Keywords: cell wall; glycoprotein; heat-stable protein

Query Match   37.1%;   Score 13;   DB 2;   Length 5;
Best Local Similarity   100.0%;   Pred. No. 2.2e+05;
Matches   3;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

```
Qy      1  TLL  3
           |||
Db      2  TLL  4
```

RESULT 3
I37027
protamine P1 - gorilla (fragment)
C;Species: Gorilla gorilla (gorilla)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
C;Accession: I37027
R;Queralt, R.; Oliva, R.
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-end
A;Reference number: I37013; MUID:94040810
A;Accession: I37027
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: EMBL:Z12145; NID:g22910; PIDN:CAA78129.1; PID:g579612

```
Query Match             37.1%;    Score 13;  DB 2;  Length 6;
Best Local Similarity   40.0%;    Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    2;  Indels    0;  Gaps    0;

Qy      2  LLEYR  6
           | ||
Db      1  MARYR  5
```

RESULT 4
S71867
glutathione transferase (EC 2.5.1.18) class alpha 5 - pig (fragment)
N;Alternate names: glutathione S-transferase class alpha 5
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 19-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 07-May-1999
C;Accession: S71867
R;Rouimi, P.; Anglade, P.; Debrauwer, L.; Tulliez, J.
Biochem. J. 317, 879-884, 1996
A;Title: Characterization of pig liver glutathione S-transferases using HPLC-electrospra
A;Reference number: S71864; MUID:96332484
A;Accession: S71867
A;Molecule type: protein
A;Residues: 1-7 <ROU>
A;Comment: At least five species-independent classes of cytosolic glutathione transferase
mitochondrial form are known.
C;Complex: dimer
C;Function:
A;Description: catalyzes the nucleophilic conjugation of intracellular glutathione to a
A;Pathway: detoxification; xenobiotics metabolism
A;Note: increased hydrophilicity of GSH-conjugates facilitates their further metabolism
es of damage
C;Superfamily: glutathione transferase
C;Keywords: dimer; transferase

```
Query Match             37.1%;    Score 13;  DB 2;  Length 7;
Best Local Similarity   50.0%;    Pred. No. 2.2e+05;
  Matches    2;  Conservative    1;  Mismatches    1;  Indels    0;  Gaps    0;

Qy      2  LLEY  5
           | ||
Db      2  ILNY  5
```

RESULT 5
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: A22565
R;Klotz, A.V.; Glazer, A.N.

J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KLO>

```
Query Match             34.3%;    Score 12;  DB 2;  Length 3;
Best Local Similarity  100.0%;    Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5  YR  6
           ||
Db      2  YR  3
```

RESULT 6
B37325
pap fimbrial regulatory protein papI - Escherichia coli
C;Species: Escherichia coli
C;Date: 11-Sep-1992 #sequence_revision 11-Sep-1992 #text_change 23-Mar-1993
C;Accession: B37325
R;Braaten, B.A.; Blyn, L.B.; Skinner, B.S.; Low, D.A.
J. Bacteriol. 173, 1789-1800, 1991
A;Title: Evidence for a methylation-blocking factor (mbf) locus involved in pap pilus
A;Reference number: A37325; MUID:91154136
A;Accession: B37325
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-5 <BRA>
A;Cross-references: GB:M63747

```
Query Match             34.3%;    Score 12;  DB 2;  Length 5;
Best Local Similarity  100.0%;    Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      4  EY  5
           ||
Db      3  EY  4
```

RESULT 7
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

```
Query Match             34.3%;    Score 12;  DB 2;  Length 5;
Best Local Similarity  100.0%;    Pred. No. 2.2e+05;
  Matches    2;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy      5  YR  6
           ||
Db      4  YR  5
```

RESULT 8
A41946
T-cell receptor gamma chain (1t.22) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: A41946

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.

GenCore version 4.5

Run on: April 21, 2002, 10:43:21 ; Search time 12.87 Seconds
            (without alignments)
            13.988 Million cell updates/sec

Title: 24
Perfect score: 24
Sequence: 1 LXXYRXXL 8

Scoring table: BLOSUM62
                 Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

OM protein   US-09-281-760c-3

Total number of hits satisfying chosen parameters: 43125

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries
                 Minimum DB seq length: 0
                 Maximum DB seq length: 8

Database: Issued_Patents_AA:*
            1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
            2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
            3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
            4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
            5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
            6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	58.3	5	2	US-08-968-676-25	Sequence 25, Appl
2	58.3	5	2	US-08-968-676-24	Sequence 24, Appl
3	58.3	7	2	US-08-968-676-22	Sequence 22, Appl
4	58.3	7	4	US-09-273-555-60	Sequence 60, Appl
5	58.3	7	4	US-09-139-802-140	Sequence 140, Appl
6	58.3	8	1	US-08-335-198-11	Sequence 11, Appl
7	58.3	8	2	US-08-968-676-21	Sequence 21, Appl
8	58.3	8	4	US-08-160-604-116	Sequence 116, Appl
9	58.3	8	4	US-08-160-604-117	Sequence 117, App
10	54.2	7	4	US-09-142-078-31	Sequence 31, Appl
11	54.2	13	4	US-09-357-141-31	Sequence 31, Appl
12	54.2	7	4	US-09-214-614-2	Sequence 2, Appl
13	50.0	3	2	US-08-968-676-27	Sequence 27, Appl
14	50.0	4	4	US-08-240-712-4	Sequence 4, Appl
15	50.0	4	1	US-08-428-488-59	Sequence 59, Appl
16	50.0	4	1	US-08-428-488-60	Sequence 60, Appl
17	50.0	4	1	US-08-428-488-61	Sequence 61, Appl
18	50.0	4	1	US-08-428-488-62	Sequence 62, Appl
19	50.0	4	1	US-08-428-488-63	Sequence 63, Appl
20	50.0	4	1	US-08-428-488-64	Sequence 64, Appl
21	50.0	4	1	US-08-428-488-65	Sequence 65, Appl
22	50.0	4	1	US-08-428-488-66	Sequence 66, Appl
23	50.0	4	1	US-08-428-488-67	Sequence 67, Appl
24	50.0	4	1	US-08-428-488-68	Sequence 68, Appl
25	50.0	4	1	US-08-428-488-69	Sequence 69, Appl
26	50.0	4	1	US-08-428-488-70	Sequence 70, Appl
27	50.0	4	1	US-08-428-488-71	Sequence 71, Appl
28	50.0	4	1	US-08-428-488-72	Sequence 72, Appl
29	50.0	4	1	US-08-428-488-73	Sequence 73, Appl
30	50.0	4	1	US-08-428-488-74	Sequence 74, Appl
31	50.0	4	2	US-08-255-272-22	Sequence 22, Appl
32	50.0	4	1	US-08-335-198-25	Sequence 25, Appl
33	50.0	4	1	US-08-429-732-20	Sequence 20, Appl
34	50.0	4	1	US-08-203-806B-15	Sequence 15, Appl
35	50.0	4	1	US-08-443-890-4	Sequence 4, Appl
36	50.0	4	1	US-08-706-178-2	Sequence 2, Appl
37	50.0	4	1	US-08-441-871-1	Sequence 1, Appl
38	50.0	4	2	US-08-441-871-63	Sequence 63, Appl
39	50.0	4	2	US-08-441-871-65	Sequence 65, Appl
40	50.0	4	2	US-08-441-871-66	Sequence 66, Appl
41	50.0	4	2	US-08-441-871-68	Sequence 68, Appl
42	50.0	4	2	US-08-441-871-70	Sequence 70, Appl
43	50.0	4	2	US-08-705-660-27	Sequence 27, Appli
44	50.0	4	2	US-08-705-660-28	Sequence 28, Appl
45	50.0	4	2	US-08-473-022-2	Sequence 2, Appli
46	50.0	4	2	US-08-623-833B-61	Sequence 61, Appl
47	50.0	4	2	US-08-685-589A-6	Sequence 6, Appli
48	50.0	4	2	US-08-685-589A-62	Sequence 62, Appli
49	50.0	4	2	US-08-968-676-26	Sequence 26, Appl
50	50.0	4	2	US-08-752-844-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1
US-08-968-676-25
Sequence 25, Application US/08968676
Patent No. 5919639
GENERAL INFORMATION:
   APPLICANT: Humphreys, Robert E
   APPLICANT: Adams, Sharlene
   APPLICANT: Xu, Minzhen
   TITLE OF INVENTION: IMMUNOTHERAPY BY MODULATION OF ANTIGEN
   NUMBER OF SEQUENCES: 165
   CORRESPONDENCE ADDRESS:
      ADDRESSEE: Kevin M. Farrell, P.C.
      STREET: P.O. Box 999
      CITY: York Harbor
      STATE: ME
      COUNTRY: USA
      ZIP: 03911
   COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: Patentin Release #1.0, Version #1.25
   CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/08/968,676
      FILING DATE:
      CLASSIFICATION: 424
   ATTORNEY/AGENT INFORMATION:
      NAME: Farrell, Kevin M
      REGISTRATION NUMBER: 35,505
      REFERENCE/DOCKET NUMBER: REH-9601
   TELECOMMUNICATION INFORMATION:
      TELEPHONE: (207) 363-0558
      TELEFAX: (207) 363-0528
INFORMATION FOR SEQ ID NO: 25:
   SEQUENCE CHARACTERISTICS:
      LENGTH: 5 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
   MOLECULE TYPE: peptide
US-08-968-676-25

Query Match   58.3%; Score 14; DB 2; Length 5;
Best Local Similarity 60.0%; Pred. No. 1.6e+05;

```
Qy      4 YRXXL 8
          | || |
Db      1 YRMKL 5

Matches   3; Conservative  0; Mismatches  2; Indels  0; Gaps  0;


RESULT 2
US-08-968-676-24
; Sequence 24, Application US/08968676
; Patent No. 5919639
; GENERAL INFORMATION:
;   APPLICANT: Humphreys, Robert E
;   APPLICANT: Adams, Sharlene
;   APPLICANT: Xu, Minzhen
;   TITLE OF INVENTION: IMMUNOTHERAPY BY MODULATION OF ANTIGEN
;   NUMBER OF SEQUENCES: 165
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Kevin M. Farrell, P.C.
;     STREET: P.O. Box 999
;     CITY: York Harbor
;     STATE: ME
;     COUNTRY: USA
;     ZIP: 03911
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/968,676
;     FILING DATE:
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Farrell, Kevin M
;     REGISTRATION NUMBER: 35,505
;     REFERENCE/DOCKET NUMBER: REH-9601
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (207) 363-0558
;     TELEFAX: (207) 363-0528
;   INFORMATION FOR SEQ ID NO: 24:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 6 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-968-676-24

Query Match        58.3%;   Score 14;  DB 2;  Length 6;
Best Local Similarity  60.0%;  Pred. No. 1.6e+05;
Matches   3; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy      4 YRXXL 8
          | || |
Db      1 YRMKL 5


RESULT 3
US-08-968-676-22
; Sequence 22, Application US/08968676
; Patent No. 5919639
; GENERAL INFORMATION:
;   APPLICANT: Humphreys, Robert E
;   APPLICANT: Adams, Sharlene
;   APPLICANT: Xu, Minzhen
;   TITLE OF INVENTION: IMMUNOTHERAPY BY MODULATION OF ANTIGEN
;   NUMBER OF SEQUENCES: 165
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE: Kevin M. Farrell, P.C.
;     STREET: P.O. Box 999
;     CITY: York Harbor
;     STATE: ME
;     COUNTRY: USA
;     ZIP: 03911
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER: IBM PC compatible
;     OPERATING SYSTEM: PC-DOS/MS-DOS
;     SOFTWARE: Patentin Release #1.0, Version #1.25
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/968,676
;     FILING DATE:
;   CLASSIFICATION: 424
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Farrell, Kevin M
;     REGISTRATION NUMBER: 35,505
;     REFERENCE/DOCKET NUMBER: REH-9601
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: (207) 363-0558
;     TELEFAX: (207) 363-0528
;   INFORMATION FOR SEQ ID NO: 22:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 7 amino acids
;     TYPE: amino acid
;     STRANDEDNESS: single
;     TOPOLOGY: linear
;     MOLECULE TYPE: peptide
US-08-968-676-22

Query Match        58.3%;   Score 14;  DB 2;  Length 7;
Best Local Similarity  60.0%;  Pred. No. 1.6e+05;
Matches   3; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy      4 YRXXL 8
          | || |
Db      1 YRMKL 5


RESULT 4
US-09-273-565-60
; Sequence 60, Application US/09273565A
; Patent No. 6166190
; GENERAL INFORMATION:
;   APPLICANT: FUJIWARA, TSUTOMU
;   APPLICANT: WATANABE, TAKESHI
;   APPLICANT: HORIE, MASATO
;   TITLE OF INVENTION: AN ISOLATED NUCLEIC ACID MOLECULE ENCODING HUMAN
;   TITLE OF INVENTION: SKELETAL MUSCLE-SPECIFIC UBIQUITIN-CONJUGATING ENZYME
;   FILE REFERENCE: Q-53599
;   CURRENT APPLICATION NUMBER: US/09/273,565A
;   CURRENT FILING DATE: 1999-03-22
;   EARLIER APPLICATION NUMBER: 09/055,699
;   EARLIER FILING DATE: 1998-04-07
;   EARLIER APPLICATION NUMBER: 08/820,170
;   EARLIER FILING DATE: 1997-03-19
;   EARLIER APPLICATION NUMBER: JP 63410/1996
;   EARLIER FILING DATE: 1996-03-19
;   EARLIER APPLICATION NUMBER: JP 69163/1997
;   EARLIER FILING DATE: 1997-03-05
;   NUMBER OF SEQ ID NOS: 95
;   SOFTWARE: Patentin Ver. 2.1
;   SEQ ID NO 60
;     LENGTH: 7
;     TYPE: PRT
;     ORGANISM: Bovine sp.
US-09-273-565-60

Query Match        58.3%;   Score 14;  DB 4;  Length 7;
Best Local Similarity  60.0%;  Pred. No. 1.6e+05;
Matches   3; Conservative  0; Mismatches  2; Indels  0; Gaps  0;

Qy      1 LXXYR 5
```

OM protein - protein search, using sw model

Copyright (c) 1993 - 2000 Compugen Ltd.
GenCore version 4.5

Run on: April 21, 2002, 10:39:31 ; Search time 26.36 Seconds
                    (without alignments)
                    22.481 Million cell updates/sec

Title: US-09-281-760C-3
Perfect score: 24
Sequence: 1 LXXYRXXL 8
Scoring table: BLOSUM62
                    Gapop 10.0 , Gapext 0.5
Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 62388

Post-processing: Minimum Match 0%
                    Maximum Match 100%
                    Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 8

Database :   A_Geneseq_1101:*
          1: /SIDS1/gcgdata/geneseq/geneseqp/AA1980.DAT:*
          2: /SIDS1/gcgdata/geneseq/geneseqp/AA1981.DAT:*
          3: /SIDS1/gcgdata/geneseq/geneseqp/AA1982.DAT:*
          4: /SIDS1/gcgdata/geneseq/geneseqp/AA1983.DAT:*
          5: /SIDS1/gcgdata/geneseq/geneseqp/AA1984.DAT:*
          6: /SIDS1/gcgdata/geneseq/geneseqp/AA1985.DAT:*
          7: /SIDS1/gcgdata/geneseq/geneseqp/AA1986.DAT:*
          8: /SIDS1/gcgdata/geneseq/geneseqp/AA1987.DAT:*
          9: /SIDS1/gcgdata/geneseq/geneseqp/AA1988.DAT:*
         10: /SIDS1/gcgdata/geneseq/geneseqp/AA1989.DAT:*
         11: /SIDS1/gcgdata/geneseq/geneseqp/AA1990.DAT:*
         12: /SIDS1/gcgdata/geneseq/geneseqp/AA1991.DAT:*
         13: /SIDS1/gcgdata/geneseq/geneseqp/AA1992.DAT:*
         14: /SIDS1/gcgdata/geneseq/geneseqp/AA1993.DAT:*
         15: /SIDS1/gcgdata/geneseq/geneseqp/AA1994.DAT:*
         16: /SIDS1/gcgdata/geneseq/geneseqp/AA1995.DAT:*
         17: /SIDS1/gcgdata/geneseq/geneseqp/AA1996.DAT:*
         18: /SIDS1/gcgdata/geneseq/geneseqp/AA1997.DAT:*
         19: /SIDS1/gcgdata/geneseq/geneseqp/AA1998.DAT:*
         20: /SIDS1/gcgdata/geneseq/geneseqp/AA1999.DAT:*
         21: /SIDS1/gcgdata/geneseq/geneseqp/AA2000.DAT:*
         22: /SIDS1/gcgdata/geneseq/geneseqp/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	16	66.7	7	20	AAY34076	Bacteroides 100 kD
2	16	66.7	7	20	AAY34080	Bacteroides 100 kD
3	16	66.7	7	20	AAY34084	Bacteroides 100 kD
4	16	66.7	8	21	AAY69583	Generic immunogeni
5	15	62.5	6	19	AAW59901	Human HLA-B27 pept
6	15	62.5	7	21	AAY51949	A. sydowi afl-SST
7	15	62.5	8	22	AAB75743	HLA class I bindin
8	14	58.3	5	2	AAP10205	Leu-enkephalin.
9	14	58.3	5	11	AAR08321	Peptide with opioi
10	14	58.3	5	15	AAR55499	Opioid oligopeptid
11	14	58.3	5	15	AAR55508	Opioid oligopeptid
12	14	58.3	5	15	AAR55490	Opioid oligopeptid
13	14	58.3	5	19	AAW47949	AE101 analogue eff
14	14	58.3	5	21	AAY69581	Generic immunogeni
15	14	58.3	5	21	AAY69582	Generic immunogeni
16	14	58.3	6	7	AAP60150	Human growth hormo
17	14	58.3	6	16	AAR83527	zif268 zinc finger
18	14	58.3	6	19	AAW50982	Growth hormone rel
19	14	58.3	6	19	AAW58999	Human HLA-B27 pept
20	14	58.3	6	19	AAW59000	Human HLA-B27 pept
21	14	58.3	6	20	AAW47948	AE101 analogue eff
22	14	58.3	6	20	AAW84426	HIV-1 nucleic acid
23	14	58.3	7	11	AAR06424	Peptide with delta
24	14	58.3	7	13	AAR24585	Immunomodulatory p
25	14	58.3	7	13	AAR24664	Immunomodulatory p
26	14	58.3	7	19	AAR60423	Tumour.homing pept
27	14	58.3	7	19	AAW47946	AE101 analogue eff
28	14	58.3	7	20	AAY42007	Rheumatoid arthrit
29	14	58.3	7	20	AAY42072	Rheumatoid arthrit
30	14	58.3	7	20	AAW93750	Mouse B16B15b mela
31	14	58.3	7	21	AAB21840	Murine melanoma ho
32	14	58.3	7	21	AAY69584	Immunogenic peptid
33	14	58.3	7	21	AAY50878	Antibody 8H.8 bind
34	14	58.3	7	22	AAU15373	Schizophrenia-asso
35	14	58.3	7	22	AAU15558	Schizophrenia-asso
36	14	58.3	7	22	AAE06418	Tumour homing pept
37	14	58.3	7	22	AAG62830	Amino acid sequenc
38	14	58.3	7	22	AAU03292	Fruit fly G protei
39	14	58.3	7	22	AAB31715	Peptide with argin
40	14	58.3	8	6	AAP51212	Analgesic peptide.
41	14	58.3	8	18	AAW24920	Vesicular stomatit
42	14	58.3	8	19	AAW47945	AE101 analogue eff
43	14	58.3	8	20	AAY46654	Immunogenic peptid
44	14	58.3	8	20	AAY47418	Immunogenic peptid
45	14	58.3	8	20	AAY10411	T cell epitope/MHC
46	14	58.3	8	21	AAY67357	Melanoma glycoprot
47	13	54.2	5	22	AAB81468	Synthetic peptide,
48	13	54.2	5	22	AAW87348	Peptide determined
49	13	54.2	6	19	AAW87354	Peptide determined
50	13	54.2	6	19	AAW87342	Peptide determined

ALIGNMENTS

RESULT 1
AAY34076

ID   AAY34076 standard; peptide; 7 AA.
XX
AC   AAY34076;
XX
DT   23-NOV-1999 (first entry)
XX
DE   Bacteroides 100 kDa antigen peptide fragment.
XX
KW   Ulcerative colitis; histone; H1-like antigen; porin antigen;
KW   Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW   perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
OS   Bacteroides caccae.
XX
PN   WO9945955-A1.
XX
PD   16-SEP-1999.
XX
PF   12-MAR-1999; 99WO-US05492.
XX
PR   12-MAR-1998; 98US-0041889.
XX
PA   (REGC) UNIV CALIFORNIA.
XX
PI   Braun J, Cohavy O;
XX
DR   WPI; 1999-551215/46.

```
RESULT   2
AAY34080
ID   AAY34080 standard; peptide; 7 AA.
XX
AC   AAY34080;
XX
DT   23-NOV-1999  (first entry)
XX
DE   Bacteroides 100 kDa antigen peptide fragment.
XX
KW   Ulcerative colitis; histone; Hl-like antigen; porin antigen;
KW   Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW   perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
XX
OS   Bacteroides caccae.
XX
PN   WO9945955-A1.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Braun J, Cohavy O;
XX
DR   WPI; 1999-551215/46.
XX
PR   12-MAR-1998;  98US-0041889.
XX
PF   12-MAR-1999;  99WO-US05492.
XX
PD   16-SEP-1999.
XX
PT   Use of histone Hl, porin or Bacteroides antigens as targets for the
PT   diagnosis, prevention and treatment of ulcerative colitis -
XX
PS   Example 6; Page 78; 134pp; English.
XX
CC   The invention provides a method for the diagnosis, prevention and
CC   treatment of ulcerative colitis (UC) using histone Hl-like antigen, a
CC   porin antigen or a Bacteroides antigen as a target antigen. The novel
CC   method of diagnosing UC in a subject suspected of having inflammatory
CC   bowel disease (IBD) comprises: (1) obtaining a sample from the subject;
CC   (2) contacting the sample with a histone Hl-like antigen, or perinuclear
CC   anti-neutrophil cytoplasmic antibody (pANCA)-reactive fragment, to form
CC   a complex of the histone Hl-like antigen, or the pANCA-reactive
CC   fragment, and antibody to the histone Hl-like antigen; and (3) detecting
CC   the presence or absence of the complex; where the presence of the
CC   complex indicates that the subject has UC. The pANCA-reactive histone
CC   Hl-like antigen, porin antigen and Bacteroides antigen are useful in the
CC   diagnosis, prevention and treatment of UC. The methods can also be used
CC   for identifying agents useful for treating UC. Sequences AAY34072-87
CC   represent peptides derived from the B. caccae 100 kDa antigen (a
CC   microbial pANCA-reactive protein).
XX
SQ   Sequence    7 AA;
XX

  Query Match             66.7%;  Score 16;  DB 20;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 4.3e+05;
  Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy          1 LXXYR 5
              | ||
Db          3 lasyr 7


RESULT   3
AAY34084
ID   AAY34084 standard; peptide; 7 AA.
XX
AC   AAY34084;
XX
DT   23-NOV-1999  (first entry)
XX
DE   Bacteroides 100 kDa antigen peptide fragment.
XX
KW   Ulcerative colitis; histone; Hl-like antigen; porin antigen;
KW   Bacteroides antigen; inflammatory bowel disease; IBD; pANCA;
KW   perinuclear anti-neutrophil cytoplasmic antibody; diagnosis.
XX
OS   Bacteroides caccae.
XX
PN   WO9945955-A1.
XX
PA   (REGC ) UNIV CALIFORNIA.
XX
PI   Braun J, Cohavy O;
XX
DR   WPI; 1999-551215/46.
XX
PR   12-MAR-1998;  98US-0041889.
XX
PF   12-MAR-1999;  99WO-US05492.
XX
PD   16-SEP-1999.
XX
PT   Use of histone Hl, porin or Bacteroides antigens as targets for the
PT   diagnosis, prevention and treatment of ulcerative colitis -
XX
PS   Example 6; Page 78; 134pp; English.
XX
CC   The invention provides a method for the diagnosis, prevention and
CC   treatment of ulcerative colitis (UC) using histone Hl-like antigen, a
CC   porin antigen or a Bacteroides antigen as a target antigen. The novel
CC   method of diagnosing UC in a subject suspected of having inflammatory
CC   bowel disease (IBD) comprises: (1) obtaining a sample from the subject;
CC   (2) contacting the sample with a histone Hl-like antigen, or perinuclear
CC   anti-neutrophil cytoplasmic antibody (pANCA)-reactive fragment, to form
CC   a complex of the histone Hl-like antigen, or the pANCA-reactive
CC   fragment, and antibody to the histone Hl-like antigen; and (3) detecting
CC   the presence or absence of the complex; where the presence of the
CC   complex indicates that the subject has UC. The pANCA-reactive histone
CC   Hl-like antigen, porin antigen and Bacteroides antigen are useful in the
CC   diagnosis, prevention and treatment of UC. The methods can also be used
CC   for identifying agents useful for treating UC. Sequences AAY34072-87
CC   represent peptides derived from the B. caccae 100 kDa antigen (a
CC   microbial pANCA-reactive protein).
XX
SQ   Sequence    7 AA;
XX

  Query Match             66.7%;  Score 16;  DB 20;  Length 7;
  Best Local Similarity   60.0%;  Pred. No. 4.3e+05;
  Matches  3;  Conservative  0;  Mismatches  2;  Indels  0;  Gaps  0;

Qy          1 LXXYR 5
              | ||
Db          3 lasyr 7
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:　　　April 21, 2002, 10:45:41 ; Search time 25.75 Seconds
　　　　　　　　(without alignments)
　　　　　　　　45.444 Million cell updates/sec

Title:　　　　US-09-281-760C-3
Perfect score:　24
Sequence:　　1 LXXYRXXL 8
Scoring table:　BLOSUM62
　　　　　　　Gapop 10.0 , Gapext 0.5
Searched:　　473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters:　352

nimum DB seq length: 0
ximum DB seq length: 8

Post-processing: Minimum Match 0%
　　　　　　　　Maximum Match 100%
　　　　　　　　Listing first 50 summaries

Database :　SPTREMBL_17:*
　　　1: sp_archea:*
　　　2: sp_bacteria:*
　　　3: sp_fungi:*
　　　4: sp_human:*
　　　5: sp_invertebrate:*
　　　6: sp_mammal:*
　　　7: sp_mhc:*
　　　8: sp_organelle:*
　　　9: sp_phage:*
　　10: sp_plant:*
　　11: sp_rodent:*
　　12: sp_virus:*
　　13: sp_vertebrate:*
　　14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	14	58.3	8	2	O52062	bacillus me
2	12	50.0	8	2	Q9R9C2	borrelia bu
3	12	50.0	8	2	P82324	borrelia bu
4	12	50.0	8	2	Q9GD47	hydriastele
5	12	50.0	8	8	Q9GD00	masoala mad
6	10	41.7	7	8	Q9GCZ4	nenga pumil
7	10	41.7	7	8	Q9I82	gnatholebia
8	9	37.5	7	8	P82618	periplaneta
9	9	37.5	7	4	O98866	spinacia ol
10	9	37.5	8	9	Q9UMH9	homo sapien
11	8	33.3	7	2	Q35792	saccharomyc
12	8	33.3	7	2	O50556	actinobacil
13	8	33.3	8	2	O09258	synechococc
14	8	33.3	8	2	Q45889	clostridium
15	8	33.3	8	2	Q9ZIE9	neisseria m
16	8	33.3	8	2	Q9RQ49	buchnera ap
17	8	33.3	8	2	O19958	gossypium b
18	8	33.3	8	8	O19956	gossypium a
19	8	33.3	8	8	O19960	gossypium m
20	8	33.3	8	8	O19961	gossypium d
21	8	33.3	8	8	Q9XNP8	boophilus m
22	8	33.3	8	8	P82324	pisum sativ
23	8	33.3	8	10	Q9ERD2	mus musculu
24	7	29.2	8	11	Q9ERD2	human adeno
25	7	29.2	7	12	Q9YVE3	human adeno
26	7	29.2	7	12	Q9YQ10	porcine tra
27	7	29.2	7	12	Q9YIR0	human adeno
28	7	29.2	7	12	Q9IQ9	human adeno
29	7	29.2	7	13	P82065	litoria rub
30	7	29.2	7	2	P72221	pseudomonas
31	7	29.2	8	2	O85406	coxiella bu
32	7	29.2	8	2	Q9S443	pseudomonas
33	7	29.2	8	2	Q9RQ57	buchnera ap
34	7	29.2	8	2	Q9R7T2	escherichia
35	7	29.2	8	2	Q9R5R2	shigella dy
36	7	29.2	8	3	P87225	saccharomyc
37	7	29.2	8	4	O15888	homo sapien
38	7	29.2	8	4	O15890	homo sapien
39	7	29.2	8	4	O15898	homo sapien
40	7	29.2	8	4	O9UJ50	homo sapien
41	7	29.2	8	4	Q9UHK1	homo sapien
42	7	29.2	8	4	Q9UDZ4	homo sapien
43	7	29.2	8	4	Q9BYY5	homo sapien
44	7	29.2	8	5	Q9TWH6	perinereis
45	7	29.2	8	6	Q9BFC3	didelphis m
46	7	29.2	8	6	Q9BFC2	macropus eu
47	7	29.2	8	6	Q9BFC1	choloepus h
48	7	29.2	8	6	Q9BFC0	choloepus d
49	7	29.2	8	6	Q9BFB9	euphractus
50	7	29.2	8	6	Q9BFB8	chaetophrac

ALIGNMENTS

RESULT 1
ID   O52062
AC   O52062;
DT   01-JUN-1998 (TrEMBLrel. 06, Created)
DT   01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT   01-AUG-1998 (TrEMBLrel. 07, Last annotation update)
DE   SIGMA FACTOR (FRAGMENT).
GN   SPOIIGB.
OS   Bacillus megaterium.
OC   Bacteria; Firmicutes; Bacillus/Clostridium group;
OC   Bacillus/Staphylococcus group; Bacillus.
OX   NCBI_TaxID=1404;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Fawcett P., Melnikov A., Youngman P.;
RL   Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF017181; AAB94056.1; -.
FT   NON_TER   8   8
SQ   SEQUENCE   8 AA;   1057 MW; F0D400441B41ADD6 CRC64;

Query Match   58.3%; Score 14; DB 2; Length 8;
Best Local Similarity   40.0%;
Matches   .2; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   1 LXXYR 5
                : ||
Db   1 MATYR 5

RESULT 2
ID   Q9R9C2
AC   Q9R9C2;   PRELIMINARY;   PRT;   8 AA.
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   PLASMID CP32-3, POSSIBLE PARTITION PROTEINS, COMPLETE CDS (FRAGMENT).
OS   Borrelia burgdorferi (Lyme disease spirochete).
OG   Plasmid cp32-3.
OC   Bacteria; Spirochaetales; Spirochaetaceae; Borrelia.
OX   NCBI_TaxID=139;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=B31;
RX   MEDLINE=98361033; PubMed=9695920;
RA   Stevenson B., Casjens S., Rosa P.;
RT   "Evidence of past recombination events among the genes encoding the
RT   Erp antigens of Borrelia burgdorferi.";
RL   Microbiology 144:0-0(0).
DR   EMBL; AF022480; AAC35445.1; -.
KW   Plasmid.
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;   985 MW;  E8B41B41A735B446 CRC64;

Qy          4 YR 5
              ||
Db          6 YR 7

Query Match             50.0%;  Score 12;  DB 2;  Length 8;
Best Local Similarity  100.0%;  Pred. No. 4.7e+05;
Matches    2;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT   3
ID   Q9GD47
AC   Q9GD47;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RIBOSOMAL PROTEIN S16 (FRAGMENT).
GN   RPS16.
OS   Hydriastele microspadix.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Arecaceae; Hydriastele.
OX   NCBI_TaxID=131271;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Asmussen C.B., Chase M.W.;
RT   "Coding and noncoding plastid DNA in palm systematics.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ404943; CAC17895.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;   975 MW;  9236C2C0441B505D CRC64;

Qy          4 YR 5
              ||
Db          3 YR 4

Query Match             50.0%;  Score 12;  DB 8;  Length 8;
Best Local Similarity  100.0%;  Pred. No. 4.7e+05;
Matches    2;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT   4
ID   Q9GD00
AC   Q9GD00;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RIBOSOMAL PROTEIN S16 (FRAGMENT).
```

```
GN   RPS16.
OS   Masoala madagascariensis.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Arecaceae; Masoala.
OX   NCBI_TaxID=131286;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Asmussen C.B., Chase M.W.;
RT   "Coding and noncoding plastid DNA in palm systematics.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ404949; CAC17925.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;   916 MW;  9D71E735A41B505D CRC64;

Qy          4 YR 5
              ||
Db          3 YR 4

Query Match             50.0%;  Score 12;  DB 8;  Length 8;
Best Local Similarity  100.0%;  Pred. No. 4.7e+05;
Matches    2;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT   5
ID   Q9GC24
AC   Q9GC24;
DT   01-MAR-2001 (TrEMBLrel. 16, Created)
DT   01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   RIBOSOMAL PROTEIN S16 (FRAGMENT).
GN   RPS16.
OS   Nenga pumila.
OG   Chloroplast.
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Arecaceae; Nenga.
OX   NCBI_TaxID=131288;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAF;
RA   Asmussen C.B., Chase M.W.;
RT   "Coding and noncoding plastid DNA in palm systematics.";
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AJ404944; CAC18505.1; -.
KW   Chloroplast.
FT   NON_TER       1       1
FT   NON_TER       8       8
SQ   SEQUENCE   8 AA;   975 MW;  9236C2C0441B505D CRC64;

Qy          4 YR 5
              ||
Db          3 YR 4

Query Match             50.0%;  Score 12;  DB 8;  Length 8;
Best Local Similarity  100.0%;  Pred. No. 4.7e+05;
Matches    2;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;


RESULT   6
ID   O99182
AC   O99182;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE   CYTOCHROME OXIDASE I (FRAGMENT).
GN   COI.
```

us-09-281-760c-3.closed.rsp

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2002, 10:46:07 ; Search time 10.34 Seconds
                                                 (without alignments)
                                                 28.367 Million cell updates/sec

Title: US-09-281-760C-3
Perfect score: 24
Sequence: 1 LXXYRXXL 8

Scoring table: BLOSUM62
                     Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 138

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Database :   SwissProt_39;*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	10	41.7	8	1	ALL6_CYDPO	P82157	cydia pomon
2	10	41.7	8	1	UPAA_HUMAN	P30096	homo sapien
3	9	37.5	6	1	OVM_LEPDE	P42985	leptinotars
4	9	37.5	7	1	HY7_PIG	P01153	sus scrofa
5	9	37.5	8	1	ALL1_CYDPO	P82152	cydia pomon
6	9	37.5	8	1	COX6_RAT	P80430	rattus norv
7	9	37.5	5	1	ALL14_CARMA	P81817	carcinus ma
8	9	33.3	6	1	FARP_MONEX	P41966	monezia ex
9	8	33.3	5	1	ALL2_CARMA	P41966	carcinus ma
10	8	33.3	7	1	ALL3_CARMA	P81805	carcinus ma
11	8	33.3	7	1	ALL4_CARMA	P81806	carcinus ma
12	8	33.3	7	1	ALL5_CARMA	P81807	carcinus ma
13	7	33.3	7	1	MNP1_LEPDE	P42984	leptinotars
14	8	33.3	8	1	ALL12_CARMA	P81815	carcinus ma
15	8	33.3	8	1	ALL15_CARMA	P81818	carcinus ma
16	8	33.3	8	1	ALL16_CARMA	P81819	carcinus ma
17	8	33.3	8	1	ALL17_CARMA	P81820	carcinus ma
18	8	33.3	8	1	ALL18_CARMA	P81821	carcinus ma
19	8	33.3	8	1	ALL13_CYDPO	P82154	cydia pomon
20	8	33.3	8	1	ALL4_CALVO	P41840	calliphora
21	8	33.3	8	1	ALL7_CARMA	P82155	cydia pomon
22	8	33.3	8	1	ALL7_CARMA	P81809	carcinus ma
23	8	33.3	8	1	ALL8_CARMA	P81811	carcinus ma
24	8	33.3	8	1	ALL9_CARMA	P81812	carcinus ma
25	8	33.3	8	1	B44K_PORGI	P81886	porphyromon
26	7	29.2	4	1	FAR3_HIRME	P42562	hirudo medi
27	7	29.2	4	1	FAR4_HIRME	P42563	hirudo medi
28	7	29.2	5	1	FARP_ARTTR	P41853	artioposthi
29	7	29.2	5	1	PRCT_PERAM	P01373	periplaneta
30	7	29.2	5	1	FARP_DAUCA	P58261	daucus caro
31	7	29.2	6	1	PSK_DAUCA	P25154	daucus caro
32	7	29.2	6	1	ACPH_RABIT	P82655	oryctolagus
33	7	29.2	7	1	ASP2_LACSN	P82655	lactobacill
34	7	29.2	7	1	ALL7_CYDPO	P82158	cydia pomon
35	7	29.2	7	1	FAR1_PROCL	P38499	procambarus
36	7	29.2	7	1	FAR2_ASCSU	P31890	ascaris suu
37	7	29.2	7	1	FAR2_PROCL	P38498	procambarus
38	7	29.2	7	1	FAR3_HAECO	P81298	haemonchus
39	7	29.2	7	1	FAR3_PANRE	P41874	panagrellus
40	7	29.2	7	1	FAR5_HIRME	P42564	hirudo medi
41	7	29.2	7	1	GFRP_MOUSE	P99025	mus musculu
42	7	29.2	7	1	UN06_PINPS	P81675	pinus pinas
43	7	29.2	7	1	WWA1_ACHFU	P35919	achatina fu
44	7	29.2	7	1	AKH_MELML	P25423	melolontha
45	8	29.2	7	1	ALL5_CALVO	P41841	calliphora
46	8	29.2	7	1	ALL5_CYDPO	P82156	cydia pomon
47	8	29.2	7	1	ANG2_BOTJA	O10582	bothrops ja
48	8	29.2	7	1	CCKN_MACEU	O30369	macropus eu
49	8	29.2	7	1	FAR3_HOMAM	P41486	homarus ame
50	8	29.2	7	1	FAR4_HOMAM	P41487	homarus ame
					LCK8_LEUMA	P19990	leucophaea

ALIGNMENTS

RESULT 1
ID   ALL6_CYDPO   STANDARD;   PRT;   8 AA.
AC   P82157;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CYDIASTATIN 6.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larva;
RX   MEDLINE=98054539; PubMed=9932829;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA   Davey M., East P.D., Thorpe A.;
RL   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES   8   8   AMIDATION.
SQ   SEQUENCE   8 AA;   936 MW;   0B2879C45B573767 CRC64;

Qy   1 LXXYRXXL 8
          | | | |
Db   1 LPLYNFGL 8

Query Match   41.7%;   Score 10;   DB 1;   Length 8;
Best Local Similarity   37.5%;   Pred. No. 1e+05;
Matches   3;   Conservative   0;   Mismatches   5;   Indels   0;   Gaps   0;

RESULT 2
ID   UPAA_HUMAN   STANDARD;   PRT;   8 AA.
AC   P30096;
DT   01-APR-1993 (Rel. 25, Created)
DT   01-APR-1993 (Rel. 25, Last sequence update)
DT   20-AUG-2001 (Rel. 40, Last annotation update)
DE   UNKNOWN PROTEIN FROM 2D-PAGE OF PLASMA (SPOT 36) (FRAGMENT).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Plasma;

```
RX   MEDLINE=93092937; PubMed=1459097;
RA   Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA   Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA   Hochstrasser D.F.;
RT   "Plasma protein map: an update by microsequencing.";
RL   Electrophoresis 13:707-714(1992).
RN   [1]
CC   -!- MISCELLANEOUS: ON THE 2D-GEL THE DETERMINED PI OF THIS UNKNOWN
CC       PROTEIN IS: 7, ITS MW IS: 12 KDA.
DR   SWISS-2DPAGE; P30096; HUMAN.
FT   NON_TER       1      1
FT   VARIANT       5      5       F -> P.
FT   NON_TER       8      8       /FTId=VAR_000004.
SQ   SEQUENCE   8 AA;  909 MW;  8667B59D1A72042 CRC64;

Query Match             41.7%;  Score 10;  DB 1;  Length 8;
Best Local Similarity   41.7%;
Matches    2;  Conservative    0;  Mismatches    2;  Indels    0;  Gaps    0;

Qy          1 LXXY 4
              |  |
Db          3 LFFY 6
```

```
RESULT 3
ID   OVM_LEPDE      STANDARD;      PRT;    6 AA.
AC   P42985;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-NOV-1995 (Rel. 32, Last annotation update)
DE   OVIDUCTAL MOTILITY STIMULATING PEPTIDE (LED-OVM).
OS   Leptinotarsa decemlineata (Colorado potato beetle).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC   Cucujiformia; Phytophaga; Chrysomeloidea; Chrysomelidae;
OC   Chrysomelinae; Leptinotarsa.
OX   NCBI_TaxID=7539;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RC   TISSUE=Head;
RX   MEDLINE=91271080; PubMed=2052497;
RA   Spittaels K., Schoofs L., Grauwels L., Smet H., van Damme J.,
RA   Proost P., Torrekens S., de Loof A.;
RT   "Isolation, identification and synthesis of novel oviductal motility
RT   stimulating head peptide in the Colorado potato beetle, Leptinotarsa
RT   decemlineata.";
RL   Peptides 12:31-36(1991).
CC   -!- FUNCTION: MYOTROPIC PEPTIDE. STIMULATES THE CONTRACTIONS OF THE
CC       OVIDUCT.
KW   Neuropeptide; Amidation.
FT   MOD_RES       6      6       AMIDATION.
SQ   SEQUENCE   6 AA;  720 MW;  6B07632B5DD03000 CRC64;

Query Match             50.0%;  Score 9;  DB 1;  Length 6;
Best Local Similarity   50.0%;
Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 YR 5
              |:|
Db          3 YK 4
```

```
RESULT 4
ID   HY7_PIG        STANDARD;      PRT;    7 AA.
AC   P01153;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   21-JUL-1986 (Rel. 01, Last annotation update)
DE   HYPOTHALAMIC HEPTAPEPTIDE.
OS   Sus scrofa (Pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX   NCBI_TaxID=9823;
RN   [1]
RP   SEQUENCE, AND SYNTHESIS.
RX   MEDLINE=81213980; PubMed=6263778;
RA   Chang R.C.C., Huang W.-Y., Arimura A., Redding T.W., Coy D.H.,
RA   Saffran M., Kong A., Hamilton J.W., Cohn D.V., Schally A.V.;
RT   "Isolation, structure and synthesis of a heptapeptide with in vitro
RT   ACTH-releasing activity from porcine hypothalamus.";
RL   Horm. Metab. Res. 13:228-232(1981).
DR   PIR; A01417; NYPG7.
SQ   SEQUENCE   7 AA;  957 MW;  632B45B1FB5059A0 CRC64;

Query Match             37.5%;  Score 9;  DB 1;  Length 7;
Best Local Similarity   37.5%;
Matches    1;  Conservative    1;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          4 YR 5
              |:|
Db          6 YK 7
```

```
RESULT 5
ID   ALL1_CYDPO     STANDARD;      PRT;    8 AA.
AC   P82152;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   30-MAY-2000 (Rel. 39, Last annotation update)
DE   CYDIASTATIN 1.
OS   Cydia pomonella (Codling moth).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
OC   Tortricoidea; Tortricidae; Olethreutinae; Cydia.
OX   NCBI_TaxID=82600;
RN   [1]
RP   SEQUENCE.
RC   TISSUE=Larva;
RX   MEDLINE=98054539; PubMed=9392829;
RA   Duve H., Johnsen A.H., Maestro J.-L., Scott A.G., Winstanley D.,
RA   Davey M., East P.D., Thorpe A.;
RT   "Lepidopteran peptides of the allatostatin superfamily.";
RL   Peptides 18:1301-1309(1997).
CC   -!- SIMILARITY: BELONGS TO THE ALLATOSTATIN FAMILY.
KW   Neuropeptide; Amidation.
FT   MOD_RES       8      8       AMIDATION.
SQ   SEQUENCE   8 AA;  934 MW;  C82B79C45B51F775 CRC64;

Query Match             37.5%;  Score 9;  DB 1;  Length 8;
Best Local Similarity   40.0%;
Matches    2;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy          4 YRXXL 8
              |   |
Db          4 YNFGL 8
```

```
RESULT 6
ID   COXG_RAT       STANDARD;      PRT;    8 AA.
AC   P80430;
DT   01-NOV-1995 (Rel. 32, Created)
DT   01-NOV-1995 (Rel. 32, Last sequence update)
DT   01-FEB-1996 (Rel. 33, Last annotation update)
DE   CYTOCHROME C OXIDASE POLYPEPTIDE VIB (EC 1.9.3.1) (AED) (FRAGMENT).
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on:　　　April 21, 2002, 10:43:41 ; Search sw model
　　　　　　　(without alignments)
　　　　　　　31.108 Million cell updates/sec

Title:　　　US-09-281-760C-3
Perfect score:　24
Sequence:　　1 LXXYRXXL 8

Scoring table:　BLOSUM62
　　　　　　　Gapop 10.0 , Gapext 0.5

Searched:　　219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters:　602

Minimum DB seq length: 0
Maximum DB seq length: 8

Post-processing: Minimum Match 0%
　　　　　　　Maximum Match 100%
　　　　　　　Listing first 50 summaries

Database :　　PIR_68:*
　　　　　1: pir1:*
　　　　　2: pir2:*
　　　　　3: pir3:*
　　　　　4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	13	54.2	6	2	I37027	protamine P1 - gor
2	12	50.0	3	2	A22565	R-phycoerythrin al
3	12	50.0	5	2	F22565	R-phycoerythrin ga
4	12	50.0	6	2	A41946	T-cell receptor ga
5	12	50.0	7	4	I56695	hypothetical L2 pr
6	12	50.0	8	2	PT0030	inulinase (EC 3.2.
7	12	50.0	6	2	S11556	hydrogensulfite re
8	9	37.5	6	2	A11490	pyruvate kinase (E
9	9	37.5	6	2	B56979	collagen alpha 1(I
10	9	37.5	7	1	NYPG7	hypothalamic hepta
11	9	37.5	8	2	S63493	dissimilatory sulf
12	9	37.5	8	2	S68802	nitrate reductase
13	9	37.5	8	2	S16324	hypothetical prote
14	9	37.5	8	2	S65381	cytochrome-c oxida
15	8	33.3	4	2	I61883	protamine P1 - ora
16	8	33.3	4	2	I37013	protamine P1 - Cer
17	8	33.3	4	2	I84419	protamine P1 - sav
18	8	33.3	5	2	B37325	pap fimbrial regul
19	8	33.3	5	2	I39964	ribosomal protein
20	8	33.3	5	2	I39966	ribosomal protein
21	8	33.3	5	2	I39965	ribosomal protein
22	8	33.3	5	2	I40469	dnaX-like protein
23	8	33.3	5	2	I49421	laminin B1 - weste
24	8	33.3	6	2	A43129	neuropeptide' GNFR
25	8	33.3	7	2	S5548	neuropeptide - Esc
26	8	33.3	8	2	B27867	mcrB protein U
27	8	33.3	8	2	D47393	neuropeptide calla
28	7	29.2	3	2	A43391	TRH-like tripeptid
29	7	29.2	3	2	S68328	blood cell protein
30	7	29.2	4	2	A32039	tyrosine-melanocyt
31	7	29.2	4	2	A37832	phenol 2-monooxyge
32	7	29.2	4	2	S09478	globulin IV alpha
33	7	29.2	4	2	PT0240	Ig heavy chain CRD
34	7	29.2	4	2	S43959	Ig mu chain V regi
35	7	29.2	5	1	HOROHA	proctolin - Americ
36	7	29.2	5	2	E60274	major protein anti
37	7	29.2	5	2	PQ0009	angiotensin conver
38	7	29.2	5	2	PQ0689	photosystem I 10.4
39	7	29.2	5	2	PQ0009	Leu-enkephalin - b
40	7	29.2	5	2	JN0861	Met-enkephalin - A
41	7	29.2	5	2	A60411	proctolin - Atlant
42	7	29.2	5	2	PT0278	Ig heavy chain CRD
43	7	29.2	5	2	S68326	blood cell protein
44	7	29.2	5	4	A58728	serrawettin W2 - S
45	7	29.2	6	2	A61411	peptidyl-dipeptida
46	7	29.2	6	2	A35890	ameletin - rat
47	7	29.2	6	2	B45510	RNA-directed DNA p
48	7	29.2	6	2	A49792	hypothetical prote
49	7	29.2	6	2	F41946	acylaminoacyl-pept
50	7	29.2	6	2	A61445	T-cell receptor ga

ALIGNMENTS

RESULT 1
I37027
protamine P1 - gorilla (fragment)
C;Species: Gorilla gorilla (gorilla)
C;Date: 04-Oct-1996 #sequence_revision 04-Oct-1996 #text_change 21-Jul-2000
R;Querait, R.; Oliva, R.
A;Accession: I37027
Gene 133, 197-204, 1993
A;Title: Identification of conserved potential regulatory sequences of the protamine-
A;Reference number: I37013; MUID:94040810
A;Accession: I37027
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-6 <RES>
A;Cross-references: EMBL:Z12145; NID:g22910; PIDN:CAA78129.1; PID:g579612

Query Match　54.2%;　Score 13;　DB 2;　Length 6;
Best Local Similarity　40.0%;　Pred. No. 2.2e+05;
Matches 2;　Conservative 1;　Mismatches 2;　Indels 0;　Gaps 0;

Qy　1 LXXYR 5
　　　: ||
Db　1 MARYR 5

RESULT 2
A22565
R-phycoerythrin alpha-1 chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: A22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: A22565
A;Molecule type: protein
A;Residues: 1-3 <KLO>

Query Match　50.0%;　Score 12;　DB 2;　Length 3;
Best Local Similarity　100.0%;　Pred. No. 2.2e+05;
Matches 2;　Conservative 0;　Mismatches 0;　Indels 0;　Gaps 0;

Qy　4 YR 5
　　　||
Qy　11

RESULT 3
F22565
R-phycoerythrin gamma-A chain - red alga (Gastroclonium coulteri) (fragment)
C;Species: Gastroclonium coulteri
C;Date: 07-Mar-1988 #sequence_revision 07-Mar-1988 #text_change 23-Mar-1993
C;Accession: F22565
R;Klotz, A.V.; Glazer, A.N.
J. Biol. Chem. 260, 4856-4863, 1985
A;Title: Characterization of the bilin attachment sites in R-phycoerythrin.
A;Reference number: A22565; MUID:85182601
A;Accession: F22565
A;Molecule type: protein
A;Residues: 1-5 <KLO>

Query Match   50.0%;   Score 12;   DB 2;   Length 5;
Best Local Similarity 100.0%;   Pred. No. 2.2e+05;
Matches   2;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     ||
Db   4 YR 5

RESULT 4
A41946
T-cell receptor gamma chain (lt.22) - mouse (fragment)
C;Species: Mus musculus (house mouse)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 07-May-1999
C;Accession: A41946
R;Whetsell, M.; Mosley, R.L.; Whetsell, L.; Schaefer, F.V.; Miller, K.S.; Klein, J.R.
Mol. Cell. Biol. 11, 5902-5909, 1991
A;Title: Rearrangement and junctional-site sequence analyses of T-cell receptor gamma ge
A;Reference number: A41946; MUID:92049316
A;Accession: A41946
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: DNA
A;Residues: 1-6 <WHE>
C;Keywords: T-cell receptor

Query Match   50.0%;   Score 12;   DB 2;   Length 6;
Best Local Similarity 100.0%;   Pred. No. 2.2e+05;
Matches   2;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     ||
Db   4 YR 5

RESULT 5
I56695
hypothetical L2 protein (mistranslated) - human papillomavirus type 16 (fragment)
C;Species: human papillomavirus type 16
C;Date: 18-Feb-2000 #sequence_revision 18-Feb-2000
C;Accession: I56695
R;Schneider-Maunoury, S.; Croissant, O.; Orth, G.
J. Virol. 61, 3295-3298, 1987
A;Title: Integration of human papillomavirus type 16 DNA sequences: a possible early eve
A;Reference number: I56695; MUID:87311896
A;Accession: I56695
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-7 <SCH>
A;Cross-references: GB:M30709; NID:g190253; PIDN:AAA65995.1; PID:g5531616
C;Comment: This is the hypothetical translation of a viral sequence integrated into the
C;Comment: It is translated in an incorrect, -1, reading frame of the L2 protein.

Query Match   50.0%;   Score 12;   DB 4;   Length 7;
Best Local Similarity 100.0%;   Pred. No. 2.2e+05;
Matches   2;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     ||
Db   1 YR 2

Query Match   50.0%;   Score 12;   DB 4;   Length 7;

RESULT 6
PT0030
inulinase (EC 3.2.1.7) - Aspergillus ficuum (fragment)
N;Alternate names: inulase
C;Species: Aspergillus ficuum
C;Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 07-May-1999
C;Accession: PT0030
R;Ettalibi, M.; Baratti, J.C.
Agric. Biol. Chem. 54, 61-68, 1990
A;Title: Molecular and kinetic properties of Aspergillus ficuum inulinases.
A;Reference number: PT0030; MUID:9034423A
A;Accession: PT0030
A;Molecule type: protein
A;Residues: 1-8 <ETT>
C;Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match   50.0%;   Score 12;   DB 2;   Length 8;
Best Local Similarity 100.0%;   Pred. No. 2.2e+05;
Matches   2;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     ||
Db   3 YR 4

RESULT 7
S11556
hydrogensulfite reductase (EC 1.8.99.3) chain 2 - Desulfovibrio thermophilus (fragment)
N;Alternate names: bisulfite reductase; desulfofuscidin
C;Species: Desulfovibrio thermophilus
C;Date: 19-Mar-1997 #sequence_revision 12-Dec-1997 #text_change 30-Jan-1998
C;Accession: S11556
R;Fauque, G.; Lino, A.R.; Czechowski, M.; Kang, L.; DerVartanian, D.V.; Moura, J.J.G.
Biochim. Biophys. Acta 1040, 112-118, 1990
A;Title: Purification and characterization of bisulfite reductase (desulfofuscidin) f
A;Reference number: S11024; MUID:90352276
A;Accession: S11556
A;Molecule type: protein
A;Residues: 1-6 <FAU>
C;Keywords: oxidoreductase

Query Match   50.0%;   Score 12;   DB 2;   Length 6;
Best Local Similarity 100.0%;   Pred. No. 2.2e+05;
Matches   2;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     ||
Db   7 YR 8

RESULT 8
A11490
pyruvate kinase (EC 2.7.1.40) - pig (fragment)
C;Species: Sus scrofa domestica (domestic pig)
C;Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 03-Mar-1995
C;Accession: A11490
R;Hjelmquist, G.; Andersson, J.; Edlund, B.; Engstrom, L.
Biochem. Biophys. Res. Commun. 61, 559-563, 1974
A;Title: Amino acid sequence of a (32-P)phosphopeptide from pig liver pyruvate kinase
A;Reference number: A11490; MUID:75127438
A;Accession: A11490
A;Molecule type: protein
A;Residues: 1-6 <HJE>

Query Match   37.5%;   Score 9;   DB 2;   Length 6;
Best Local Similarity 50.0%;   Pred. No. 2.2e+05;
Matches   1;   Conservative   1;   Mismatches   0;   Indels   0;   Gaps   0;

Qy   4 YR 5
     |:
Db   5 YK 6

OM protein - protein search, using sw model

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

Run on: April 21, 2002, 10:15:59 ; Search time 94.43 Seconds
            2.621 Million cell updates/sec
            (without alignments)

Title: US-09-281-760C-5
Perfect score: 63
Sequence: 1 GMNLTWYRESK 11
Scoring table: BLOSUM62
                Gapop 10.0 , Gapext 0.5
Searched: 212252 seqs, 22503292 residues
Total number of hits satisfying chosen parameters: 212252

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 50 summaries

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Database :   Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Match Length	DB	ID	Description
1	63	100.0	426	1	US-08-336-583-2	Sequence 2, Appli
2	63	100.0	426	5	PCT-US95-13795-2	Sequence 2, Appli
3	41	65.1	583	2	US-08-432-016-2	Sequence 2, Appli
4	41	65.1	583	2	US-08-684-594-2	Sequence 2, Appli
5	38	60.3	107	4	US-09-240-274-46	Sequence 46, Appl
6	38	60.3	256	4	US-08-484-905-113	Sequence 119, App
7	38	60.3	256	3	US-08-481-985B-113	Sequence 113, App
8	38	60.3	256	4	US-08-370-476-113	Sequence 113, App
9	38	60.3	500	4	US-08-960-190A-25	Sequence 25, Appl
10	37	58.7	500	4	US-08-117-362-3	Sequence 3, Appl
11	37	58.7	860	1	US-08-486-924-3	Sequence 3, Appl
12	37	58.7	860	1	US-08-117-362-4	Sequence 4, Appli
13	37	58.7	894	1	US-08-486-924-4	Sequence 4, Appli
14	36	57.1	894	1	US-08-276-852-113	Sequence 113, App
15	36	57.1	96	1	US-08-899-575-113	Sequence 113, App
16	36	57.1	96	1	PCT-US95-08743-113	Sequence 113, App
17	36	57.1	96	5	US-08-569-147-74	Sequence 74, Appl
18	36	57.1	128	4	US-08-417-495-25	Sequence 25, Appl
19	36	57.1	182	2	US-08-284-391B-25	Sequence 25, Appl
20	36	57.1	182	4	US-09-218-950-25	Sequence 25, Appl
21	36	57.1	182	4	PCT-US92-01785-25	Sequence 25, Appl
22	36	57.1	182	5	PCT-US95-00454-25	Sequence 25, Appl
23	36	57.1	182	5	5457037-3	Patent No. 5457037
24	36	57.1	750	6	5457037-5	Patent No. 5457037
25	36	57.1	751	6	US-08-410-784A-5	Sequence 5, Appli
26	36	57.1	772	2	US-09-346-237-4	Sequence 4, Appli
27	36	57.1	776	4	US-09-346-237-7	Sequence 7, Appli
28	36	57.1	776	4	US-09-346-237-29	Sequence 29, Appl
29	35	55.6	776	4	US-08-497-312-29	Sequence 31, Appl
30	35	55.6	110	1	US-08-497-312-31	Sequence 34, Appl
31	35	55.6	110	1	US-09-113-750A-34	Sequence 119, App
32	35	55.6	246	4	US-08-413-118-119	Sequence 119, App
33	35	55.6	319	3	US-08-473-446-119	Sequence 2, Appli
34	35	55.6	391	1	PCT-US95-15696-2	Sequence 12, Appl
35	35	55.6	459	1	US-08-220-151-12	Sequence 14, Appl
36	35	55.6	459	1	US-08-220-151-14	Sequence 12, Appl
37	35	55.6	459	1	US-08-413-118-12	Sequence 14, Appl
38	35	55.6	459	3	US-08-473-446-12	Sequence 12, Appl
39	35	55.6	459	3	US-09-213-053-6	Sequence 14, Appl
40	35	55.6	459	4	US-08-473-446-14	Sequence 14, Appl
41	35	55.6	459	4	US-09-213-053-6	Sequence 4, Appli
42	35	55.6	661	2	US-09-514-014-4	Sequence 4, Appli
43	35	55.6	661	2	US-08-833-823-4	Sequence 13, Appl
44	35	55.6	668	1	US-08-232-538-13	Sequence 13, Appl
45	35	55.6	668	2	US-08-786-164-13	Sequence 3, Appl
46	35	55.6	714	2	US-08-990-114-3	Sequence 2, Appli
47	35	55.6	767	2	US-08-874-678-2	Sequence 2, Appli
48	35	55.6	767	1	US-08-643-839-2	Sequence 15, Appl
49	35	55.6	788	1	US-08-232-538-15	Sequence 15, Appl
50	35	55.6	788	2	US-08-786-164-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1
US-08-336-583-2
Sequence 2, Application US/08335583
Patent No. 5629415

GENERAL INFORMATION:
APPLICANT: HOLLIS, GREGORY F.
APPLICANT: PATEL, MAYUR D.
TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULIN E
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTINE E. CARTY
STREET: 126 E. LINCOLN AVENUE
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0900
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/336,583
FILING DATE: 09-NOV-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: CARTY, CHRISTINE E.
REGISTRATION NUMBER: 36,099
REFERENCE/DOCKET NUMBER: 19211
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-6734
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-336-583-2

Query Match   100.0%; Score 63; DB 1; Length 426;
Best Local Similarity   100.0%; Pred. No. 0.0014;
Matches   11; Conservative   0; Mismatches   0; Indels   0; Gaps   0;

L X YR

```
Qy     1 GMNLTWYRESK  11
           |||||||||||
Db   245 GMNLTWYRESK 255
```

RESULT 2
PCT-US95-13795-2
Sequence 2, Application PC/TUS9513795
GENERAL INFORMATION:
APPLICANT: HOLLIS, GREGORY F.
APPLICANT: PATEL, MAYUR D.
NUMBER OF SEQUENCES: 4
TITLE OF INVENTION: DNA ENCODING CANINE IMMUNOGLOBULINS
CORRESPONDENCE ADDRESS:
ADDRESSEE: CHRISTINE E. CARTY
STREET: 126 E. LINCOLN AVENUE; P.O. BOX 2000
CITY: RAHWAY
STATE: NEW JERSEY
COUNTRY: USA
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/13795
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: CARTY, CHRISTINE E.
REGISTRATION NUMBER: 36,099
REFERENCE/DOCKET NUMBER: 19211Y
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-6734
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 426 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

Query Match 100.0%; Score 63; DB 5; Length 426;
Best Local Similarity 100.0%; Pred. No. 0.0014;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

PCT-US95-13795-2

```
Qy     1 GMNLTWYRESK  11
           |||||||||||
Db   245 GMNLTWYRESK 255
```

RESULT 3
US-08-432-016-2
Sequence 2, Application US/08432016
Patent No. 5968768
GENERAL INFORMATION:
APPLICANT: HAYNES, BARTON F.
APPLICANT: ARUFFO, ALEJANDRO
APPLICANT: PATEL, DHAVALKUMAR
APPLICANT: BOWEN, MICHAEL A.
APPLICANT: MARQUARDT, HANS
TITLE OF INVENTION: CD6 LIGAND
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/432,016
FILING DATE: 01-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/333,350
FILING DATE: 02-NOV-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/143,903
FILING DATE: 02-NOV-1993
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 1579-95
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 583 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
FEATURE:
OTHER INFORMATION:
OTHER INFORMATION: Xaa = M/S
OTHER INFORMATION: Xaa = M/T
US-08-432-016-2

```
Qy     3 NLTWYRESK  11
           |:|||| |
Db   167 NITWYRNGK 175
```

RESULT 4
US-08-684-594-2
Sequence 2, Application US/08684594
Patent No. 5998172
GENERAL INFORMATION:
APPLICANT: HAYNES, BARTON F.
APPLICANT: ARUFFO, ALEJANDRO
APPLICANT: PATEL, DHAVALKUMAR
APPLICANT: BOWEN, MICHAEL A.
APPLICANT: MARQUARDT, HANS
TITLE OF INVENTION: CD6 LIGAND
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/684,594
FILING DATE: 18-JUL-1996

Query Match 65.1%; Score 41; DB 2; Length 583;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/432,016
; FILING DATE: 01-MAY-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/333,350
; FILING DATE: 02-NOV-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/143,903
; FILING DATE: 02-NOV-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 1579-112
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 583 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; OTHER INFORMATION: Xaa = M/S
; OTHER INFORMATION: Xaa = M/T
US-08-684-594-2

Query Match   65.1%; Score 41; DB 2; Length 583;
Best Local Similarity 66.7%; Pred. No. 15;
Matches   6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy   3 NLTWYRESK 11
         |:|||
Db 167 NITWYRNGK 175

RESULT 5
US-09-240-274-46
Sequence 46, Application US/09240274
Patent No. 6255455
GENERAL INFORMATION:
; APPLICANT: Siegel, Donald L.
; TITLE OF INVENTION: Rh(D)-BINDING PROTEINS AND MAGNETICALLY ACTIVATED CELL
; TITLE OF INVENTION: SORTING METHOD FOR PRODUCTION THEREOF
; FILE REFERENCE: 0959G-42U2
; CURRENT APPLICATION NUMBER: US/09/240,274
; CURRENT FILING DATE: 1999-01-29
; EARLIER APPLICATION NUMBER: 60/081,380
; EARLIER FILING DATE: 1998-04-10
; EARLIER APPLICATION NUMBER: 60/028,550
; EARLIER FILING DATE: 1996-10-11
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 46
; LENGTH: 107
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: anti-Rh(D) chain II6

Query Match   60.3%; Score 38; DB 4; Length 107;
Best Local Similarity 50.0%; Pred. No. 7.5;
Matches   5; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy   1 GMNLTWYRES 10
         | ||| |
Db 29 GENLNWYQQT 38

RESULT 6
US-08-484-905-113
Sequence 113, Application US/08484905
Patent No. 5976551
GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS-/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 113:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 256 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-113

Query Match   60.3%; Score 38; DB 2; Length 256;
Best Local Similarity 60.0%; Pred. No. 20;
Matches   6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy   2 MNLTWYRESK 11
         :|:|| | |
Db 144 INITWLRNSK 153

RESULT 7
US-08-481-985B-113
Sequence 113, Application US/08481985B
Patent No. 6011146
GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Phillipe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
ADDRESSEE: Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/481,985B
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03495.0106-04000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-481-985B-113

Qy   2 MNLTWYRESK 11
        :|:||-||||-||
Db 144 INITWLRNSK 153

Query Match   60.3%;   Score 38; DB 3; Length 256;
Best Local Similarity 60.0%;   Pred. No. 20;
Matches   6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 8
US-08-370-476-113
Sequence 113, Application US/08370476
Patent No. 6153408
GENERAL INFORMATION:
APPLICANT: Mottez, Estelle
APPLICANT: Abastado, Jean-Pierre
APPLICANT: Kourilsky, Phillipe
APPLICANT: Lone, Yu-Chun
APPLICANT: Ojcius, David
APPLICANT: Castrouge, Armanda
TITLE OF INVENTION: Altered Major Histocompatibility Complex
NUMBER OF SEQUENCES: 127
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/370,476
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,575
FILING DATE: 07-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/072,787
FILING DATE: 06-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/801,818
FILING DATE: 05-DEC-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/792,473
FILING DATE: 15-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 05243.0001-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4000
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 256 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-370-476-113

Qy   2 MNLTWYRESK 11
        :|:||-||||-||
Db 144 INITWLRNSK 153

Query Match   60.3%;   Score 38; DB 4; Length 256;
Best Local Similarity 60.0%;   Pred. No. 20;
Matches   6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

RESULT 9
US-08-960-190A-25
Sequence 25, Application US/08960190A
Patent No. 6232445
GENERAL INFORMATION:
APPLICANT: Rhode, Peter R.
APPLICANT: Acevedo, Jorge
APPLICANT: Burkhardt, Martin
APPLICANT: Jiao, Jin-an
APPLICANT: Wong, Hing C.
TITLE OF INVENTION: SOLUBLE MHC COMPLEXES AND
TITLE OF INVENTION: METHODS OF USE THEREOF
NUMBER OF SEQUENCES: 38
CORRESPONDENCE ADDRESS:
ADDRESSEE: Dike, Bronstein, Roberts & Cushman, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: usa
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/960,190A
FILING DATE: 29-OCT-1997
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:

```
;   ATTORNEY/AGENT INFORMATION:
;     NAME: Corless, Peter F
;     REGISTRATION NUMBER: 33,860
;     REFERENCE/DOCKET NUMBER: 48002
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE: 617-523-3400
;     TELEFAX: 617-523-6440
;     TELEX:
;
US-08-960-190A-25
; FRAGMENT TYPE: Internal
  MOLECULE TYPE: protein
;     TOPOLOGY: linear
;     STRANDEDNESS: single
;     TYPE: amino acid
;     LENGTH: 500 amino acids
; SEQUENCE CHARACTERISTICS:
INFORMATION FOR SEQ ID NO: 25:

Db      388  INITWLRNSK  397
            :|:||||:||
Qy        2  MNLTWYRESK  11

Matches    6; Conservative    2; Mismatches    2; Indels    0; Gaps    0;
Best Local Similarity   60.0%; Pred. No. 41;
Query Match             60.3%;  Score 38; DB 4; Length 500;


RESULT 10
US-08-117-362-3
Sequence 3, Application US/08117362
Patent No. 5595872
GENERAL INFORMATION:
  APPLICANT: Wetterau II, John R.
  APPLICANT: Sharp, Daru Y.
  APPLICANT: Gregg, Richard E.
  TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
  NUMBER OF SEQUENCES: 32
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Burton Rodney
    STREET: P.O. Box 4000
    CITY: Princeton
    STATE: New Jersey
    COUNTRY: U.S.A.
    ZIP: 08543-4000
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/117,362
    FILING DATE: 03-SEP-1993
    CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Gaul, Timothy J.
    REGISTRATION NUMBER: 33,111
    REFERENCE/DOCKET NUMBER: DC21b
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (609) 252-5901
    TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 860 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-117-362-3

Query Match             58.7%;  Score 37; DB 1; Length 860;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
```

```
Matches    6; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy        6  WYRESK  11
             ||||||
Db      736  WYRESK  741


RESULT 11
US-08-486-924-3
Sequence 3, Application US/08486924
Patent No. 5789197
GENERAL INFORMATION:
  APPLICANT: Wetterau II, John R.
  APPLICANT: Sharp, Daru Y.
  APPLICANT: Gregg, Richard E.
  TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
  NUMBER OF SEQUENCES: 32
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Burton Rodney
    STREET: P.O. Box 4000
    CITY: Princeton
    STATE: New Jersey
    COUNTRY: U.S.A.
    ZIP: 08543-4000
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/486,924
    FILING DATE: 07-JUN-1995
    CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER: US 08/117,362
    FILING DATE: 03-SEP-1993
  ATTORNEY/AGENT INFORMATION:
    NAME: Gaul, Timothy J.
    REGISTRATION NUMBER: 33,111
    REFERENCE/DOCKET NUMBER: DC21b
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (609) 252-5901
    TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
  LENGTH: 860 amino acids
  TYPE: amino acid
  STRANDEDNESS: single
  TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-486-924-3

Query Match             58.7%;  Score 37; DB 1; Length 860;
Best Local Similarity  100.0%;  Pred. No. 1.1e+02;
Matches    6; Conservative    0; Mismatches    0; Indels    0; Gaps    0;

Qy        6  WYRESK  11
             ||||||
Db      736  WYRESK  741


RESULT 12
US-08-117-362-4
Sequence 4, Application US/08117362
Patent No. 5595872
GENERAL INFORMATION:
  APPLICANT: Wetterau II, John R.
  APPLICANT: Sharp, Daru Y.
  APPLICANT: Gregg, Richard E.
  TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
  NUMBER OF SEQUENCES: 32
  CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Burton Rodney
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/117,362
FILING DATE: 03-SEP-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J.
REGISTRATION NUMBER: 33,111
REFERENCE/DOCKET NUMBER: DC21b
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 252-5901
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-117-362-4

Qy   6 WYRESK 11
      |||||
Db   770 WYRESK 775

Query Match   58.7%;   Score 37;   DB 1;   Length 894;
Best Local Similarity   100.0%;   Pred. No. 1.2e+02;
Matches   6;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 13
US-08-486-924-4
Sequence 4, Application US/08486924
Patent No. 5789197
GENERAL INFORMATION:
APPLICANT: Wetterau II, John R.
APPLICANT: Sharp, Daru Y.
APPLICANT: Gregg, Richard E.
TITLE OF INVENTION: MICROSOMAL TRIGLYCERIDE TRANSFER PROTEIN
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Burton Rodney
STREET: P.O. Box 4000
CITY: Princeton
STATE: New Jersey
COUNTRY: U.S.A.
ZIP: 08543-4000
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/486,924
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/117,362
FILING DATE: 03-SEP-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gaul, Timothy J.
REGISTRATION NUMBER: 33,111

REFERENCE/DOCKET NUMBER: DC21b
TELECOMMUNICATION INFORMATION:
TELEPHONE: (609) 252-5901
TELEFAX: (609) 252-4526
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 894 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-486-924-4

Qy   6 WYRESK 11
      |||||
Db   770 WYRESK 775

Query Match   58.7%;   Score 37;   DB 1;   Length 894;
Best Local Similarity   100.0%;   Pred. No. 1.2e+02;
Matches   6;   Conservative   0;   Mismatches   0;   Indels   0;   Gaps   0;

RESULT 14
US-08-276-852-113
Sequence 113, Application US/08276852
Patent No. 5652138
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
Patent Counsel
STREET: 10666 No. 5652138th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/276,852
FILING DATE: 18-JUL-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCR1452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-276-852-113

Query Match
Best Local Similarity   57.1%;   Score 36;   DB 1;   Length 96;
Matches   5;   Conservative   55.6%;   Pred. No. 15;
          2;   Mismatches   2;   Indels   0;   Gaps   0;

Qy   1 GMNLTWYRE   9
        | ||| ||::
Db 17 GTNLAWYQQ 25

RESULT 15
US-08-899-575-113
Sequence 113, Application US/08899575
Patent No. 5770440
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5770440th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
PRIOR APPLICATION DATA:
CLASSIFICATION: 435
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-899-575-113
Sequence 113, Application US/08899575
Patent No. 5804440
GENERAL INFORMATION:
APPLICANT: Burton, Dennis R
APPLICANT: Barbas, Carlos F
APPLICANT: Lerner, Richard A
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
CORRESPONDENCE ADDRESS:
ADDRESSEE: The Scripps Research Institute, Office of
ADDRESSEE: Patent Counsel
STREET: 10666 No. 5804440th Torrey Pines Road, Suite 220,
STREET: Mail Drop TPC8
CITY: La Jolla
STATE: CA
COUNTRY: USA
ZIP: 92037
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/899,575
FILING DATE: 24-JUL-1997
PRIOR APPLICATION DATA:
CLASSIFICATION:
APPLICATION NUMBER: US 08/276,852
FILING DATE: 18-JUL-1994
APPLICATION NUMBER: US 08/178,302
FILING DATE: 30-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/954,148
FILING DATE: 30-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: SCRI452P
TELECOMMUNICATION INFORMATION:
TELEPHONE: 619-554-2937
TELEFAX: 619-554-6312
INFORMATION FOR SEQ ID NO: 113:
SEQUENCE CHARACTERISTICS:
LENGTH: 96 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-899-575-113

Qy   1 GMNLTWYRE   9
        | ||| ||::
Db 17 GTNLAWYQQ 25

RESULT 17
PCT-US95-08743-113
Sequence 113, Application PCT/US9508743
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: HUMAN NEUTRALIZING MONOCLONAL ANTIBODIES
TITLE OF INVENTION: TO HUMAN IMMUNODEFICIENCY VIRUS
NUMBER OF SEQUENCES: 170
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

Query Match
Best Local Similarity   57.1%;   Score 36;   DB 1;   Length 96;
Matches   5;   Conservative   55.6%;   Pred. No. 15;
          2;   Mismatches   2;   Indels   0;   Gaps   0;

RESULT 16
US-08-899-575-113

Qy   1 GMNLTWYRE   9
        | ||| ||::
Db 17 GTNLAWYQQ 25

Query Match
Best Local Similarity   57.1%;   Score 36;   DB 1;   Length 96;
Matches   5;   Conservative   55.6%;   Pred. No. 15;
          2;   Mismatches   2;   Indels   0;   Gaps   0;

```
SOFTWARE:        PatentIn Release #1.0, Version #1.25 (EPO)
CURRENT APPLICATION DATA:
  APPLICATION NUMBER:    PCT/US95/08743
  FILING DATE:           11-JUL-1995
PRIOR APPLICATION DATA:
  APPLICATION NUMBER:    US 08/276,852
  FILING DATE:           18-JUL-1994
INFORMATION FOR SEQ ID NO: 113:
  SEQUENCE CHARACTERISTICS:
    LENGTH:      96 amino acids
    TYPE:        amino acid
    TOPOLOGY:    linear
  MOLECULE TYPE: protein
PCT-US95-08743-113
```

```
Query Match      57.1%;  Score 36;  DB 5;  Length 96;
Best Local Similarity   55.6%;  Pred. No. 15;
Matches   5;  Conservative   2;  Mismatches   2;  Indels   0;  Gaps   0;

Qy   1 GMNLTWYRE 9
         | ||:||::
Db  17 GTNLAWYQQ 25
```

```
RESULT 18
US-08-569-147-74
Sequence 74, Application US/08569147
Patent No. 6180377
GENERAL INFORMATION:
  APPLICANT:
  TITLE OF INVENTION:  HUMANISED ANTIBODIES
  NUMBER OF SEQUENCES:  95
  CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Woodcock Washburn Kurtz Mackiewicz &
    ADDRESSEE:  No. 61803771iS, LLP
    STREET:     One Liberty Place - 46th Floor
    CITY:       Philadelphia
    STATE:      PA
    COUNTRY:    U.S.A.
    ZIP:        19103
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  Floppy disk
    COMPUTER:     IBM PC compatible
    OPERATING SYSTEM:  PC-DOS/MS-DOS
    SOFTWARE:     Patentin Release #1.0, Version #1.25 (EPO)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/569,147
    FILING DATE:  25-March-1996
    CLASSIFICATION:  536
  ATTORNEY/AGENT INFORMATION:
    NAME:  Trujillo, Doreen Yatko
    REGISTRATION NUMBER:  35,719
    REFERENCE/DOCKET NUMBER:  CARP-0047
  TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (215) 568-3100
    TELEFAX:    (215) 568-3439
INFORMATION FOR SEQ ID NO: 74:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  128 amino acids
    TYPE:    amino acid
    TOPOLOGY:  linear
  MOLECULE TYPE:  protein
US-08-569-147-74
```

```
Query Match      57.1%;  Score 36;  DB 4;  Length 128;
Best Local Similarity   55.6%;  Pred. No. 20;
Matches   5;  Conservative   2;  Mismatches   2;  Indels   0;  Gaps   0;

Qy   3 NLTWYRESK 11
         ||:|||::|
Db  52 NLAWYROKQ 60
```

```
RESULT 19
US-08-417-495-25
Sequence 25, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
  APPLICANT:  Seed, Brian et al.
  TITLE OF INVENTION:  Redirection of Cellular Immunity by Chimeras
  TITLE OF INVENTION:  Receptor
  NUMBER OF SEQUENCES:  27
  CORRESPONDENCE ADDRESS:
    ADDRESSEE:  Fish & Richardson
    STREET:     225 Franklin Street
    CITY:       Boston
    STATE:      MA
    COUNTRY:    USA
    ZIP:        02110-2804
  COMPUTER READABLE FORM:
    MEDIUM TYPE:  3.5" Diskette, 1.44 Mb
    COMPUTER:     IBM PS/2 Model 50z or 55SX
    OPERATING SYSTEM:  IBM P.C. DOS (Version 3.30)
    SOFTWARE:     Wordperfect (Version 5.0)
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER:  US/08/417,495
    FILING DATE:
    CLASSIFICATION:  435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:  US/08/203,866
    FILING DATE:
    APPLICATION NUMBER:  US/07/847,566
    FILING DATE:
    APPLICATION NUMBER:  07/665,961
    FILING DATE:  March 7, 1991
    CLASSIFICATION:  435
  ATTORNEY/AGENT INFORMATION:
    NAME:  Clark, Paul T.
    REGISTRATION NUMBER:  30,162
    REFERENCE/DOCKET NUMBER:  00786/119002
  TELECOMMUNICATION INFORMATION:
    TELEPHONE:  (617) 542-5070
    TELEFAX:    (617) 542-8906
    TELEX:      200154
INFORMATION FOR SEQ ID NO: 25:
  SEQUENCE CHARACTERISTICS:
    LENGTH:  182 amino acids
    TYPE:    amino acid
    TOPOLOGY:  linear
  MOLECULE TYPE:  amino acid
US-08-417-495-25
```

```
Query Match      57.1%;  Score 36;  DB 2;  Length 182;
Best Local Similarity   44.4%;  Pred. No. 30;
Matches   4;  Conservative   4;  Mismatches   1;  Indels   0;  Gaps   0;

Qy   3 NLTWYRESK 11
         | :||:::|
Db  52 NITWFKDGK 60
```

```
RESULT 20
US-08-284-391B-25
Sequence 25, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
  APPLICANT:  Seed, Brian
  APPLICANT:  Banapour, Babak
  APPLICANT:  Romeo, Charles
  APPLICANT:  Kolanus, Waldemar
  TITLE OF INVENTION:  TARGETED CYTOLYSIS OF HIV-INFECTED
  TITLE OF INVENTION:  CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
  NUMBER OF SEQUENCES:  53
```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
ZIP: 02110
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-284-391B-25

Query Match   57.1%;   Score 36;   DB 2;   Length 182;
Best Local Similarity   44.4%;   Pred. No. 30;
Matches   4; Conservative   4; Mismatches   1; Indels   0; Gaps   0;

Qy   3 NLTWYRESK 11
         |:||:||::|
Db   52 NITWFKDGK 60

ULT 21
09-218-950-25
Sequence 25, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
ZIP: 02110
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS

SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-218-950-25

Query Match   57.1%;   Score 36;   DB 4;   Length 182;
Best Local Similarity   44.4%;   Pred. No. 30;
Matches   4; Conservative   4; Mismatches   1; Indels   0; Gaps   0;

Qy   3 NLTWYRESK 11
         |:||:||::|
Db   52 NITWFKDGK 60

RESULT 22
PCT-US92-01785-25
Sequence 25, Application PC/TUS9201785
GENERAL INFORMATION:
APPLICANT: The General Hospital Corporation
TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor
TITLE OF INVENTION: Chimeras
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
ZIP: 02110-2804
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/01785
FILING DATE: 19920306
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:

```
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: amino acid
PCT-US92-01785-25
```

```
Qy    3  NLTWYRESK 11
         |:||:::|
b    52  NITWFKDGK 60
```

```
Query Match       57.1%;  Score 36;  DB 5;  Length 182;
Best Local Similarity  44.4%;  Pred. No. 30;
Matches    4;  Conservative  4;  Mismatches  1;  Indels  0;  Gaps  0;
```

```
RESULT 23
PCT-US95-00454-25
Sequence 25, Application PC/TUS9500454
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected
TITLE OF INVENTION: Cells by Chimeric CD4 Receptor-
TITLE OF INVENTION: Bearing Cells
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/00454
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/847,566
FILING DATE: March 6, 1992
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 182 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: amino acid
PCT-US95-00454-25
```

```
Query Match       57.1%;  Score 36;  DB 5;  Length 182;
Best Local Similarity  44.4%;  Pred. No. 30;
```

```
Matches    4;  Conservative  4;  Mismatches  1;  Indels  0;  Gaps  0;

Qy    3  NLTWYRESK 11
         |:||:::|
Db   52  NITWFKDGK 60

RESULT 24
5457037-3
Patent No. 5457037
APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI GUIDO;DI GENNARO, CARLO
TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
NUMBER OF SEQUENCES: 7
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/1,797
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 749,621
FILING DATE: 19-AUG-1991
APPLICATION NUMBER: 224,114
FILING DATE: 25-JUL-1988
SEQ ID NO:3:
LENGTH: 750
5457037-3
```

```
Query Match       57.1%;  Score 36;  DB 6;  Length 750;
Best Local Similarity  60.0%;  Pred. No. 1.4e+02;
Matches    6;  Conservative  1;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1  GMNLTWYRES 10
         | ||||:|
Db  640  GSQLTWYQPS 649

RESULT 25
5457037-5
Patent No. 5457037
APPLICANT: TOGNONI, ANGELO;CARRERA, PAOLO;CAMERINI, BARBARA;
GALLI, GIULIANO;LUCCHESE, GIUSEPPE;GRANDI GUIDO;DI GENNARO, CARLO
TITLE OF INVENTION: CLONING OF THE GENE CODING THE ISOAMYLASE
ENZYME AND ITS USE IN THE PRODUCTION OF SAID ENZYME
NUMBER OF SEQUENCES: 7
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/1,797
FILING DATE: 08-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 749,621
FILING DATE: 19-AUG-1991
APPLICATION NUMBER: 224,114
FILING DATE: 25-JUL-1988
SEQ ID NO:5:
LENGTH: 751
5457037-5
```

```
Query Match       57.1%;  Score 36;  DB 6;  Length 751;
Best Local Similarity  60.0%;  Pred. No. 1.4e+02;
Matches    6;  Conservative  1;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1  GMNLTWYRES 10
         | ||:|
Db  641  GSQLTWYQPS 650

Search completed: April 21, 2002, 10:16:00
Job time: 285 sec
```